



Final Program and Abstracts



Eighth International Congress on Peer Review and Scientific Publication

Enhancing the quality and credibility of science

September 10-12, 2017

Swissôtel

Chicago, Illinois USA

peerreviewcongress.org



Eighth International Congress on Peer Review and Scientific Publication

Table of Contents

Welcome	2
Advisory Board	3
Program Highlights	4
Satellite Sessions	4
Program	5
Sunday, September 10	5
Monday, September 11	6
Tuesday, September 12	8
Poster Abstract Titles	10
Plenary Session Abstracts	16
Poster Session Abstracts	53
Sponsors	118
Exhibitors	120
Congress Organizers and Planners	123

Eighth International Congress on Peer Review and Scientific Publication

Enhancing the quality and credibility of science

Welcome!

JAMA, the *BMJ*, and METRICS welcome you to Chicago and the Eighth International Congress on Peer Review and Scientific Publication. Our aim is to encourage research with the quality and credibility of peer review and scientific publication, to establish the evidence base on which scientists can improve the conduct, reporting, and dissemination of scientific research. We have continued our efforts to broaden the scope of the Congress to all aspects of peer review and publication—from funding to postpublication—and to all sciences.

We will have 3 days for presentations of new research into peer review and all aspects of scientific publication: day 1, common problems; day 2, improvements; and day 3, innovations.

There are 45 research plenary session presentations. Each plenary session presentation will be followed by equal time for discussion and questions from the audience. In addition, there are 84 research poster presentations scheduled for Monday and Tuesday, and 5 plenary session Invited Talks.

We hope you will take an active part in the program, as we depend on your participation in the discussion sessions to make the Congress a success. Enjoy the Congress and enjoy Chicago!

Drummond Rennie, Congress Director

Annette Flanagan, Congress Executive Director

Fiona Godlee, European Director

Theodora Bloom, European Coordinator

Michael Berkwits, Associate Director

Steve Goodman, Associate Director

John P.A. Ioannidis, Associate Director



Follow us on Twitter @peerrevcongress.



Engage in conversations via Twitter and Facebook using #PRC80.

The plenary sessions will be streamed live on social media.

Peer Review Congress Advisory Board

Congress Director

Drummond Rennie

University of California, San Francisco
San Francisco, California, USA

Congress Executive Director

Annette Flanagan

JAMA and The JAMA Network
Chicago, Illinois, USA

European Director

Fiona Godlee

BMJ
London, UK

European Coordinator

Theodora Bloom

BMJ
London, UK

Associate Directors

Michael Berkwits

JAMA and The JAMA Network
Chicago, Illinois, USA

Steve Goodman

Meta-Research Innovation Center at
Stanford (METRICS)
Stanford, California, USA

John P.A. Ioannidis

Meta-Research Innovation Center at
Stanford (METRICS)
Stanford, California, USA

Members

Douglas G. Altman

University of Oxford
Oxford, UK

Vivienne Bachelet

Medwave and Universidad de Santiago
de Chile
Santiago, Chile

Ginny Barbour

Griffith University School of Medicine
Brisbane, Queensland, Australia

Howard Bauchner

JAMA and The JAMA Network
Chicago, Illinois, USA

Jeffrey Beall

University of Colorado, Denver
Denver, Colorado, USA

Lisa A. Bero

The University of Sydney
Sydney, NSW, Australia

Isabelle Boutron

French Cochrane Center
Paris, France

Edward W. Campion

New England Journal of Medicine
Boston, Massachusetts, USA

Phil Davis

Phil Davis Consulting
Ithaca, New York, USA

Kay Dickersin

Johns Hopkins Bloomberg School of
Public Health
Baltimore, Maryland, USA

Dan Evanko

American Association for Cancer
Research
Philadelphia, Pennsylvania, USA

Peter C. Gøtzsche

Nordic Cochrane Center
Copenhagen, Denmark

Trish Groves

BMJ
London, UK

R. Brian Haynes

McMaster University Faculty of Health
Sciences
Hamilton, Ontario, Canada

Véronique Kiermer

PLOS
San Francisco, California, USA

Sabine Kleinert

The Lancet
London, UK

Christine Laine

Annals of Internal Medicine
Philadelphia, Pennsylvania, USA

José Florencio F. Lapeña Jr

University of the Philippines
Manila, Philippines

Emilie Marcus

Cell
Cambridge, Massachusetts, USA

Ana Marušić

Journal of Global Health
Zagreb, Croatia

David Moher

Ottawa Hospital Research Institute
Ottawa, Ontario, Canada

David Ofori-Adjei

Ghana Medical Journal
Accra, Ghana

Jigisha Patel

BioMed Central, Springer Nature
London, UK

Peush Sahni

The National Medical Journal of India
New Delhi, India

David Schriger

University of California, Los Angeles
Los Angeles, California, USA

Richard Smith

icddr,b
Dhaka, Bangladesh

Martin R. Tramèr

European Journal of Anaesthesiology
Geneva, Switzerland

Valda Vinson

Science
Washington DC, USA

Erik Von Elm

Cochrane Switzerland
Lausanne, Switzerland

Li Youping

Journal of Evidence-Based Medicine
Chengdu, China

Deborah A. Zarin

National Library of Medicine
Bethesda, Maryland, USA

Program Highlights

Three days of Original Research

September 10

Common Problems in Peer Review and Scientific Publication

September 11

Improving Peer Review and Scientific Publication

September 12

Innovations in Peer Review and Scientific Publication

45 Plenary Session reports of original research

84 Poster Session reports of original research

Equal time for presentation and audience participation

Plenary Session Invited Talks

Peer Review: Are We Looking at the Right Things?

Lisa A. Bero (Australia)

Social Dynamics and Structural Bias in Peer Review, 1865-1965

Aileen Fyfe (United Kingdom)

Custodians of High-Quality Science: Are Editors and Peer Reviewers Good Enough?

David Moher (Canada)

Statistical Review in Biomedical Journals: Can We Get SMARTA?

Steve N. Goodman (United States)

Preprints and Other Threats to Traditional Publishing

Harlan M. Krumholz (United States)

Satellite Sessions

Saturday, September 9

8:00 AM - 12:00 PM

EQUATOR GoodReports Campaign Workshop
Implementing Reporting Guidelines: Time for Action

For more information:

www.equator-network.org/2017/04/07/workshop_prc8_chicago_2017/

1:30 PM - 5:00 PM

EQUATOR and Centre of Journalology Workshop
What Should Journals Do in the Wake of Predatory Journals?

For more information:

ohri.ca/journalology/events.aspx/

Monday, September 11

7:00 AM - 7:45 AM

WAME Business Meeting

5:30 PM

8th Annual EQUATOR Lecture

What a Reporting Guideline Can Do: 15 Years of STARD
(Standards for Reporting Diagnostic Accuracy Studies)

Patrick M. Bossuyt

For more information: www.equator-network.org/2017/01/04/equator-lecture-2017-patrick-bossuyt/

Tuesday, September 12

5:30 PM - 7:00 PM

Peer Review Week

Under the Microscope: Transparency in Peer Review

For more information: peerreviewweek.wordpress.com/activities/

Eighth International Congress on Peer Review and Scientific Publication

September 10–12, 2017

Program

All plenary sessions will be held in the Zurich Ballroom, D-F. Poster sessions will be held in the St Gallen and Montreux rooms.

Breaks and Exhibits will be held in Zurich Foyer and Zurich A-C. Luncheons will be served in the Vevey room.

Sunday, September 10, 2017

Common Problems in Peer Review and Scientific Publication

7:00 AM - 8:00 AM

**Registration, Continental Breakfast, and
Visit Exhibits**

8:00 AM

Welcome

Drummond Rennie (United States)

8:05 AM

**Peer Review: Are We Looking at the Right
Things?**

Lisa A. Bero (Australia)

Moderator: Drummond Rennie

8:30 AM

**Bias Associated With Conflict of Interest and
Peer Review**

Moderator: Véronique Kiermer (United States)

**The Prevalence of Conflict of Interest
Disclosures in Biomedical Research**

Quinn Grundy, Adam Dunn, Florence Bourgeois, and Lisa
A. Bero (Australia, United States)

**The Influence of Industry Funding and
Other Financial Conflicts of Interest on the
Outcomes and Quality of Systematic Reviews**

Camilla Hansen, Andreas Lundh, Kristine Rasmussen, Tove
Faber Frandsen, Peter C. Gøtzsche, and Asbjørn Hróbjartsson
(Denmark)

**Analysis of Uptake and Outcome in Author-
selected Single-blind vs Double-blind Peer
Review at *Nature Journals***

Elisa De Ranieri, Barbara McGillivray, Sowmya
Swaminathan, Michelle Samarasinghe, and Leah Gruen
(United Kingdom, United States)

**Gender and Age Bias in Peer Review in Earth
and Space Science Journals**

Jory Lerback and R. Brooks Hanson (United States)

9:50 AM

Refreshment Break and Visit Exhibits

10:30 AM

**Bias in Reporting and Publication of
Research**

Moderator: Martin R. Tramèr (Switzerland)

**Augmenting Systematic Reviews With
Information From ClinicalTrials.gov to
Increase Transparency and Reduce Bias**

Stacy Springs, **Gaelan Adam**, Thomas Trikalinos, John W.
Williams Jr, Jennifer Eaton, Megan Von Isenburg, Jennifer
M. Gierisch, Lisa M. Wilson, Ritu Sharma, Sydney M. Dy,
Julie M. Waldfogel, Karen A. Robinson, Meera Viswanathan,
Jennifer Cook Middleton, Valerie L. Forman-Hoffman, Elise
Berliner, and Robert M. Kaplan (United States)

**Bias Associated With Publication of Interim
Results of Randomized Trials: A Systematic
Review**

Steven Woloshin, Lisa M. Schwartz, Pamela J. Bagley,
Heather B. Blunt, and Brian White (United States)

Identification and Classification of Spin in Clinical Studies Evaluating Biomarkers in Ovarian Cancer: A Systematic Review

Mona Ghannad, Maria Olsen, and Patrick M. Bossuyt
(France, the Netherlands)

Spin in Published Biomedical Literature: A Systematic Review

Quinn Grundy, Kellia Chiu, and Lisa A. Bero (Australia)

12:00 PM - 1:30 PM

Lunch and Visit Exhibits

1:30 PM

Integrity and Misconduct

Moderator: Jigisha Patel (United Kingdom)

Summary Effect Sizes in Meta-analyses After Removal of Retracted Studies From the Pool of Primary Studies

Daniele Fanelli and David Moher (Canada, United States)

Assessing the Outcomes of Introducing a Digital Image Quality Control Review Into the Publication Process for Research Articles in Physiology Journals

Rita Scheman and **Christina N. Bennett** (United States)

Fact Checking Nucleotide Sequences in Life Science Publications: The Seek & Blastn Tool

Jennifer A. Byrne and **Cyril Labbé** (Australia, France)

Types of Research Integrity Issues Encountered by a Specialist Research Integrity Group

Magdalena Morawska and **Stephanie L. Boughton** (United Kingdom)

2:50 PM

Refreshment Break and Visit Exhibits

3:30 PM

Data Sharing

Moderator: Howard Bauchner (United States)

Early Experiences With Journal Data Sharing Policies: A Survey of Published Clinical Trial Investigators

Sara Tannenbaum, Joseph S. Ross, Harlan M. Krumholz, Nihar R. Desai, Jessica D. Ritchie, Richard Lehman, Ginger M. Gamble, Jacqueline Bachand, Sara Schroter, Trish Groves, and Cary P. Gross (United Kingdom, United States)

Sharing Data Through the Yale University Open Data Access (YODA) Project: Early Experience

Joseph S. Ross, Jessica D. Ritchie, Stephen Bamford, Jesse A. Berlin, Karla Childers, Nihar Desai, Ginger M. Gamble, Cary P. Gross, Richard S. Lehman, Peter Lins, Sandra A. Morris, Joanne Waldstreicher, and Harlan M. Krumholz (United Kingdom, United States)

Statements About Intent to Share Individual Participant Data at ClinicalTrials.gov

Annice Bergeris, Tony Tse, **Deborah A. Zarin** (United States)

4:30 PM

Social Dynamics and Structural Bias in Peer Review, 1865-1965

Aileen Fyfe (United Kingdom)

Moderator: Drummond Rennie

5:00 PM

Welcome Reception

Monday, September 11, 2017

Improving Peer Review and Scientific Publication

7:00 AM - 8:00 AM

Registration, Continental Breakfast, and Visit Exhibits

7:00 AM - 7:45 AM

WAME Business Meeting

8:00 AM

Morning Welcome and Housekeeping

Michael Berkwits (United States)

8:05 AM

Custodians of High-Quality Science: Are Editors and Peer Reviewers Good Enough?

David Moher (Canada)

Moderator: Michael Berkwits

8:30 AM

Quality of Reporting

Moderator: Douglas G. Altman (United Kingdom)

Association of Journal-Level and Study-Level Variables With Proximity of Primary Study Results to Summary Estimates From Meta-Analyses in Imaging Journals

Robert Frank, Matthew McInnes, Deborah Levine, Herbert Kressel, Julie Jesurum, William Petrreich, Trevor McGrath, and Patrick M. Bossuyt (Canada, the Netherlands, United States)

Discrepancies in Reporting Between Trial Publications and Clinical Trial Registries in High-Impact Journals

Sarah Daisy Kosa, Lawrence Mbuagbaw, Victoria Borg Debono, Mohit Bhandari, Brittany B. Dennis, Gabrielle Ene, Alvin Leenus, Daniel Shi, Michael Thabane, Thuva Vanniyasingam, Chenglin Ye, Elgene Yranon, Shiyuan Zhang, and Lehana Thabane (Canada)

Methodological and Reporting Quality of Systematic Reviews Underpinning Clinical Practice Guidelines

Cole Wayant and Matt Vassar (United States)

Optimism Bias in Contemporary National Clinical Trial Network Phase 3 Trials

Kaveh Zakeri, Sonal S. Noticewala, Lucas K. Vitzthum, Elena Sojourner, and Loren K. Mell (United States)

9:50 AM

Refreshment Break and Visit Exhibits

10:20 AM

Quality of the Scientific Literature

Moderator: Ana Marušić (Croatia)

Scientific Quality in a Series of Comparative Effectiveness Research Studies

Harold Sox, Evan Mayo-Wilson, Kelly Vander Ley, Marina Broitman, David Hickam, Steven Clauser, Yen-Pin Chiang, and Evelyn Whitlock (United States)

Pitfalls in the Use of Statistical Methods in Systematic Reviews of Therapeutic Interventions: A Cross-sectional Study

Matthew J. Page, Douglas G. Altman, Larissa Shamseer, Joanne E. McKenzie, Nadera Ahmadzai, Dianna Wolfe, Fatemeh Yazdi, Ferrán Catalá-López, Andrea C. Tricco, and David Moher (Australia, Canada, Spain, United Kingdom)

Introducing Reporting Guidelines and Checklists for Contributors to *Radiology*: Results of an Author and Reviewer Survey

Marc Dewey, Deborah Levine, Patrick M. Bossuyt, and Herbert Y. Kressel (Germany, the Netherlands, United States)

Reported Use of Standard Reporting Guidelines Among *JNCI* Authors, Editorial Outcome, and Reviewer Ratings Related to Adherence to Guidelines and Clarity of Presentation

Jeannine Botos (United States)

Impact of an Intervention to Improve Compliance With the ARRIVE Guidelines for the Reporting of In Vivo Animal Research

Emily Sena for the Intervention to Improve Compliance With the ARRIVE Guidelines (IICARus) Collaborative Group (United Kingdom)

12:00 PM - 1:30 PM

Lunch and Visit Exhibits

1:30 PM

Trial Registration

Moderator: Kay Dickersin (United States)

Association of Trial Registration With Reporting of Clinical Trials: Comparison of Protocols, Registries, and Published Articles

An-Wen Chan, Annukka Pello, Jessica Kitchen, Anna Axentiev, Jorma Virtanen, Annie Liu, and Elina Hemminki (Canada, Finland)

Impact of FDAAA on Registration, Results Reporting, and Publication of Neuropsychiatric Clinical Trials Supporting FDA New Drug Approval, 2005-2014

Constance X. Zou, Jessica E. Becker, Adam T. Phillips, Harlan M. Krumholz, Jennifer E. Miller, and Joseph S. Ross (United States)

Evaluation of the ClinicalTrials.gov Results Database and Its Relationship to the Peer-Reviewed Literature

Deborah A. Zarin, Tony Tse, **Rebecca J. Williams**, Thiyaagu Rajakannan, and Kevin M. Fain (United States)

2:30 PM - 3:30 PM

Poster Sessions

Refreshments and Visit Exhibits

3:40 PM

Funding/Grant Review

Moderator: Trish Groves (United Kingdom)

Geographic and Gender Bias in Peer Review of Applications Submitted to the Swiss National Science Foundation

João Martins, François Delavy, Anne Jorstad, and Matthias Egger (Switzerland)

Stakeholder Perceptions of Peer Review at the National Institutes of Health Center for Scientific Review

Mary Ann Guadagno and Richard K. Nakamura (United States)

Testing of 2 Application Ranking Approaches at the National Institutes of Health Center for Scientific Review

Richard K. Nakamura, Amy L. Rubinstein, Adrian P. Vancea, and Mary Ann Guadagno (United States)

Scientist, Patient, and Stakeholder Roles in Research Application Review: Analysis of the Patient-Centered Outcomes Research Institute (PCORI) Approach to Research Funding

Laura P. Forsythe, Lori B. Frank, Tsahai A. Tafari, Sarah S. Cohen, Michael Lauer, Steve Clauser, Christine Goertz, and Suzanne Schrandt (United States)

5:00 PM

Statistical Review in Biomedical Journals: Can We Get SMARTA?

Steve N. Goodman (United States)

Moderator: Christine Laine

5:30 PM

8th Annual EQUATOR Lecture

Tuesday, September 12, 2017

Innovations in Peer Review and Scientific Publication

7:00 AM - 8:00 AM

Registration, Continental Breakfast, and Visit Exhibits

8:00 AM

Morning Welcome and Housekeeping

John P. A. Ioannidis (United States)

8:05 AM

Preprints and Other Threats to Traditional Publishing

Harlan M. Krumholz (United States)

Moderator: John P. A. Ioannidis

8:30 AM

Peer Review Innovations

Moderator: Fiona Godlee (United Kingdom)

Assessment of Author Demand for Double-blind Peer Review in IOP (Institute of Physics) Publishing Journals

Simon Harris, Marc Gillett, Pernille Hammelsoe, and Tim Smith (United Kingdom)

Use of Open Review by Discipline, Country, and Over Time: An Analysis of Reviews and Journal Policies Posted on Publons

Sarah Parks, Salil Gunashekar, and Elta Smith (United Kingdom)

Comparison of Acceptance of Peer Reviewer Invitations by Peer Review Model: Open, Single-blind, and Double-blind Peer Review

Maria Kowalczyk and Michelle Samarasinghe (United Kingdom, United States)

A Novel Open Peer Review Format for an Emergency Medicine Blog

Scott Kobner, Derek Sifford, and Michelle Lin (United States)

9:50 AM

Refreshment Break and Visit Exhibits

10:30 AM

Editorial and Peer-Review Process Innovations

Moderator: David Schriger (United States)

Impact of a Change in Editorial Policy at the Nature Publishing Group (NPG) on Their Reporting of Biomedical Research

Malcolm Macleod for the NPQIP Collaborative Group (United Kingdom)

Assessment of Signing Peer Reviews in Principle and in Practice at Public Library of Science (PLOS) Journals

Elizabeth Seiver and Helen Atkins (United States)

The Role of Persistent Identifiers in the Peer Review Process: Use of ORCID

Alice Meadows (United States)

Introduction of Patient Review Alongside Traditional Peer Review at a General Medical Journal (*The BMJ*): A Mixed Methods Study

Sara Schroter, Amy Price, Rosamund Snow, Tessa Richards, Sam Parker, Elizabeth Loder, and Fiona Godlee (United Kingdom, United States)

12:00 PM - 1:30 PM

Lunch and Visit Exhibits

1:30 PM

Prepublication and Postpublication Issues

Moderator: Theo Bloom (United Kingdom)

Associations Between bioRxiv Preprint NonCitation Attention and Publication in the Biomedical Literature: A Cross-sectional and Cohort Study

Stylianos Serghiou and John P. A. Ioannidis (United States)

Differences in Readership Metrics and Media Coverage Among Negative, Positive, or Mixed Studies Published by the *New England Journal of Medicine*

Ramya Ramaswami, Sagar Deshpande, Rebecca Berger, Pamela Miller, and Edward W. Campion (United States)

Reproducible Research Practices in Systematic Reviews of Therapeutic Interventions: A Cross-sectional Study

Matthew J. Page, Douglas G. Altman, Larissa Shamseer, Joanne E. McKenzie, Nadera Ahmadzai, Dianna Wolfe, Fatemeh Yazdi, Ferrán Catalá-López, Andrea Tricco, and David Moher (Australia, Canada, Spain, United Kingdom)

2:30 PM - 3:30 PM

Poster Sessions

Refreshments and Visit Exhibits

3:40 PM

Postpublication Issues

Moderator: Emilie Marcus (United States)

Analysis of Indexing Practices of Corrected and Republished Articles in MEDLINE, Web of Science, and Scopus

Tea Marasović, Ana Utrobičić, and **Ana Marušić** (Croatia)

A Cross-sectional Study of Commenters and Commenting in PubMed, 2014-2016: Who's Who in PubMed Commons

Melissa D. Vaught, Diana C. Jordan, and Hilda Bastian (United States)

The Role of PubPeer Comments in Alerting Editors to Serious Problems With Clinical Research Publications

Elizabeth Wager and Emma Veitch (United Kingdom)

4:40 PM

Closing Session

A Tribute to Drummond Rennie

Annette Flanagan, Fiona Godlee, and Howard Bauchner (United Kingdom, United States)

5:00 PM

Adjournment

5:30 PM - 7:00 PM

Post-Congress Satellite Panel: Peer Review Week

Under the Microscope: Transparency in Peer Review

Poster Session Abstracts

Posters will be presented during 1 of 2 sessions, on Monday, September 11, and Tuesday, September 12. Specific day of presentation is listed before each abstract title.

Authorship and Contributorship

MONDAY

Trends in Authorship and Team Science in Major Medical Journals, 2005-2015

Christopher C. Muth and Robert M. Golub (United States)

TUESDAY

Frequency of Reporting on Patient Involvement in Research Studies Published in a General Medical Journal: A Descriptive Study

Amy Price, Sara Schroter, Rosamund Snow, Sophie Staniszewska, Sam Parker, and Tessa Richards (United Kingdom)

MONDAY

Authorship for Sale: A Survey Among Predatory Publishers and Journals

Pravin M. Bolshete (India)

MONDAY

A Survey of Awareness of Authorship Criteria by Clinical Investigators and Medical Writers in China

Jing-ling Bao, Xiu-yuan Hao, Wei-zhu Liu, Pei-fang Wei, Yang Pan, Jun-min Wei, and Young-mao Jiang (China)

MONDAY

Survey of Authors' Views on Barriers to Preparation of Biomedical Research Manuscripts

June Oshiro, Suzanne L. Caubet, Kelly Viola, and Jill M. Huber (United States)

MONDAY

Researchers' Awareness and Use of Authorship Guidelines: An International Survey

Sara Schroter, Ilaria Montagni, Elizabeth Loder, Matthias Eikermann, Elke Schaeffner, and Tobias Kurth (France, Germany, United Kingdom, United States)

TUESDAY

International Survey of Researchers' Experiences With and Attitudes Toward Coauthorship in the Humanities and Social Sciences

Tiffany Drake, Bruce Macfarlane, and Mark Robinson (United Kingdom)

Bias in Peer Review, Reporting, and Publication

TUESDAY

Financial Ties and Discordance Between Results and Conclusions in Trials of Weight Loss and Physical Activity Apps

Veronica Yank, Sanjhavi Agarwal, Rhea Red, and Amy Lozano (United States)

TUESDAY

Bias Arising From the Use of Patient-Reported Outcome Measures

Joel J. Gagnier, Jianyu Lai, and Chris Robbins (United States)

MONDAY

Discrepancies in Risk of Bias Judgments for Randomized Trials of Acupuncture Included in More Than 1 Cochrane Review

Yonggang Zhang, Linli Zheng, Youping Li, Mike Clarke, and Liang Du (China, United Kingdom)

MONDAY

Gender Bias in Funding of Proposals Submitted to the Swiss National Science Foundation

François Delavy, Anne Jorstad, and Matthias Egger (Switzerland)

TUESDAY

Prevalence of High or Unclear Risk of Bias Assessments in Diagnostic Accuracy Studies Included in Cochrane Reviews

Nicola Di Girolamo, Reint Meursinge Reynders, and Alexandra Winter (Italy, Netherlands, United States)

TUESDAY

Assessment of Agreement Between Reviewers in the Open Postpublication Peer Review Process of *F1000Research*
Tiago Barros and Liz Allen (United Kingdom)

MONDAY

An Update on Reporting Bias in the Antidepressant Literature: An FDA-Controlled Examination of Drug Efficacy
Erick H. Turner, Sepideh Alavi, Andrea Cipriani, Toshi Furukawa, Ilya Ivlev, Ryan McKenna, and Yusuke Ogawa (Japan, United Kingdom, United States)

MONDAY

Prevalence of Comparative Effectiveness Trials of Surgical vs Medical Interventions
Anaïs Rameau, Anirudh Saraswathula, Ewoud Schuit, and John P. A. Ioannidis (the Netherlands, United States)

TUESDAY

Frequency of Citation of Clinical Trials With a High Risk of Bias
Priyanka Desai, Mary Butler, and Robert L. Kane (United States)

Bibliometrics and Scientometrics

TUESDAY

The Clinical Impact of Published Trials
Ashwini R. Sehgal (United States)

MONDAY

Association Between the Journal Evaluation Program of the Korean Association of Medical Journal Editors (KAMJE) and Change in Quality of Member Journals
Hee-Jin Yang, Se Jeong Oh, and Sung-Tae Hong (Korea)

MONDAY

Association of Publication Rate With the Award of Starting and Advanced Grants
David Pina, Lana Barać, Ivan Buljan, and Ana Marušić (Belgium, Croatia)

TUESDAY

Determining the Appropriateness of *Pediatrics* Case Reports Citations
Bryan A. Sisk, Griffin S. Collins, Claire Dillenbeck, and J. Jeffrey Malatack (United States)

Conflict of Interest

MONDAY

Reporting of Conflicts of Interest of Panel Members Formulating Clinical Practice Guidelines in Anesthesiology: A Cross-sectional Study
Damien Wyssa, Martin R. Tramèr, and Nadia Elia (Switzerland)

TUESDAY

Physician Journal Editors and the Open Payments Program
Victoria S. S. Wong, Lauro Nathaniel Avalos, and Michael L. Callahan (United States)

MONDAY

Effect of Different Financial Competing Interest Statements on Readers' Perceptions of Clinical Educational Articles: A Randomized Controlled Trial
Sara Schroter, Julia Pakpoor, Julie Morris, Mabel Chew, and Fiona Godlee (United Kingdom)

TUESDAY

Competing Interest Disclosures Compared With Industry Payments Reporting Among Highly Cited Authors in Clinical Medicine
Daniel M. Cook and Kyle Kaminski (United States)

MONDAY

Collaboration Between Industry and Academics in Clinical Vaccine, Drug, and Device Trials: A Survey of Academic Investigators
Kristine Rasmussen, Lisa A. Bero, Rita Redberg, Peter C. Göttsche, and Andreas Lundh (Australia, Denmark, United States)

TUESDAY

Accuracy, Transparency, and Conflict of Interest in Medical Journal Drug Advertisements
James R. Scott, Mark Gibson, and Rebecca S. Benner (United States)

Data Sharing

MONDAY

Data Sharing Policies in Scholarly Publications: Interdisciplinary Comparisons
Michal Tal-Socher and Adrian Ziderman (Israel)

TUESDAY

Endorsement of Data Sharing by Authors and High-Impact Medical Journals in China: A Survey of Authors and Assessment of Journal Online Instructions

Yuanyuan Ji, Limin Chen, Xiuyuan Hao, Ningning Wang, and Yalin Bao (China)

Dissemination of Information

TUESDAY

NEJM Quick Take Videos: A Survey of Authors and Readers

Rebecca Berger, Ramya Ramaswami, Karen Buckley, Roger Feinstein, Kathy Stern, Timothy Vining, Stephen Morrissey, and Edward W. Campion (United States)

MONDAY

Age of Clinical Trial Data at the Time of Publication: A Systematic Review of Clinical Trials Published in 2015

John W. Welsh, Yuan Lu, Sanket S. Dhruva, Behnood Bikdeli, Nihar R. Desai, Liliya Benchetrit, Chloe O. Zimmerman, Lin Mu, Joseph S. Ross, and Harlan M. Krumholz (United States)

TUESDAY

Publication and Dissemination of Results of Clinical Trials in Neurology

Anirudh Sreekrishnan, **David Mampre**, Cora Ormseth, Laura Miyares, Audrey Leasure, Lindsay Klickstein, Joseph S. Ross, and Kevin N. Sheth (United States)

MONDAY

Disclosure of Results of Clinical Trials Sponsored by Pharmaceutical Companies

Slavka Baronikova, Jim Purvis, Christopher Winchester, Eric Southam, Julie Beeso, and Antonia Panayi (Switzerland, United Kingdom)

TUESDAY

Frequency and Format of Clinical Trial Results Disseminated to Participants: A Survey of Trialists

Sara Schroter, Amy Price, Mario Malički, Rosamund Snow, Tessa Richards, and Mike Clarke (Croatia, United Kingdom)

Editorial and Peer Review Process

MONDAY

Editorial Rejections in *Obstetrics & Gynecology*

Randi Y. Zung, Rebecca S. Benner, and Nancy C. Chescheir (United States)

TUESDAY

Implementation of a Peer Reviewer Probation and Performance Monitoring Program at *The Journal of Bone & Joint Surgery*

Marc Swiontkowski and **Christina Nelson** (United States)

TUESDAY

Evaluation of a Software Tool and Process to Assist Authors With Reporting Ethical Requirements of Studies

Tracy Ronan and Alice Ellingham (United Kingdom)

TUESDAY

Peer Review in Croatian Open Access Scholarly and Professional Journals: A Cross-Disciplinary Survey

Ivana Hebrang Grgić and Jadranka Stojanovski (Croatia)

MONDAY

Feasibility of a Randomized Controlled Trial Comparing Results-Blind Peer Review vs Standard Peer Review for Reducing Publication Bias in an Open Peer Review Journal

Katherine S. Button, Anna Clark, Tim Shipley, and Liz Bal (United Kingdom)

MONDAY

A Scoping Review of the Roles and Tasks of Peer Reviewers in the Biomedical Journal Editorial Process

Ketevan Glonti, Daniel Cauchi, Erik Cobo, Isabelle Boutron, David Moher, Darko Hren (Canada, Croatia, France, Malta, Spain, United States)

MONDAY

A Survey of Chinese Medical Researchers' Awareness and Use of Author-Suggested Reviewers

Limin Chen, Xiuyuan Hao, **Yuanyuan Ji**, and Yalin Bao (China)

Funding/Grant Peer Review

MONDAY

Influences of Independent Peer Reviewer Scores on UK National Institute for Health Research Grant Funding Recommendations

Nicola McArdle, **Helen Payne**, Sheila Turner, and Jeremy C. Wyatt (United Kingdom)

TUESDAY

Influence of Evaluation Criteria on Overall Assessment in Peer Review of Project Grants Submitted to the Swiss National Science Foundation

Stéphanie Würth, Katrin Milzow, and Matthias Egger (Switzerland)

TUESDAY

Patient-Centered Outcomes Research Institute (PCORI) Methodology Standards to Improve the Design and Reporting of Research

Evan Mayo-Wilson, Kelly Vander Ley, Kay Dickersin, and Mark Helfand (United States)

Integrity and Misconduct

TUESDAY

The Journal Project of the Russian Dissernet

Andrey Rostovtsev, Alexei Kassian, **Vasiliy Vlassov**, Anna Abalkina, and Larisa Melikhova (Russia)

MONDAY

Misuse of Received Manuscripts by Peer Reviewers: A Cross-sectional Survey

Darren Taichman, Jill Jackson, Deborah Cotton, Cynthia Mulrow, Jaya Rao, Mary Beth Schaeffer, Catharine Stack, Sankey Williams, and Christine Laine (United States)

MONDAY

A Survey of Knowledge and Perception of Plagiarism Among Chinese Authors and Reviewers

Pei-Fang Wei, Xiu-Yuan Hao, Yang Pan, Wei-Zhu Liu, Jing-Ling Bao, Jun-Min Wei, and Yong-Mao Jiang (China)

TUESDAY

Post-retraction Citations in Korean Medical Journals

Sun Huh, **Hyun Jung Yi**, Hye-Min Cho, and Soo Young Kim (Korea)

TUESDAY

Assessment of a Standardized Tool to Identify Deceptive Journals

Kathleen Berryman, Sheree Crosby, Lacey Earle, and Lucas Toutloff (United States)

MONDAY

Assessment of the Prevalence of Integrity Issues in Submitted Manuscripts

Damian Pattinson and Chrissy Prater (United States)

Peer Review

TUESDAY

A Comparison of Reviewer Contribution Distributions in Publons

Andrew Preston and Tom Culley (United Kingdom)

MONDAY

A Pilot Study of Online Training of Patient Stakeholders Aimed at Improving Knowledge and Skills to Complete Peer Review of PCORI Draft Final Research Reports

Karen B. Eden, Ilya Ivlev, Amy Forester, Camber Hansen-Karr, Ed Reid, Lauren Saxton, Kelly Vander Ley, and Mark Helfand (United States)

MONDAY

Identification of Motivations for Peer Reviewers to Perform Prepublication Review of Manuscripts: A Systematic Review

Mersiha Mahmić-Kaknjo, Mario Malički, Ana Utrobičić, Dario Sambunjak, and Ana Marušić (Bosnia and Herzegovina, Croatia)

TUESDAY

Publishing Peer Review Reports Alongside Articles With Separate DOIs: A Pilot Study of 5 Journals in Different Scientific Disciplines

Bahar Mehmani (the Netherlands)

TUESDAY

Peer Review in Computer Science Conferences Published by Springer

Mario Malički, Martin Mihajlov, Aliaksandr Birukou, and Volha Bryl (Croatia, Germany, Macedonia)

MONDAY

An Analysis of Peer Review Cases Brought to the Committee on Publication Ethics (COPE) From 1997-2016

Elizabeth C. Moylan, Virginia Barbour, Linda Gough, Charon A. Pierson, Deborah Poff, Natalie Ridgeway, Michael Wise, and Adrian Ziderman (Australia, Canada, Israel, United Kingdom, United States)

MONDAY

Modeling the Effects of Jointly Implemented Peer Review Systems on Scientific Publication

Michail Kovanis, Philippe Ravaut, and Raphael Porcher (France)

TUESDAY

Linguistic Features in Peer Reviewer Reports: How Peer Reviewers Communicate Their Recommendations

Ketevan Glonti, Darko Hren, Simon Carter, and Sara Schroter (Croatia, France, United Kingdom)

MONDAY

Assessment of Regional Diversity of Reviewers in Journals Published in Medicine and Agricultural and Biological Sciences

Thomas Gaston and Pippa Smart (United Kingdom)

Quality of Reporting

TUESDAY

Characterizing Major Issues in ClinicalTrials.gov Results Submissions

Heather D. Dobbins, Cassiah Cox, Tony Tse, Rebecca J. Williams, and Deborah A. Zarin (United States)

MONDAY

Transparency in Cross-National Research: Quality of Reporting

Elena Damian, Bart Meuleman, and Wim van Oorschot (Belgium)

TUESDAY

Presence of a Unique Trial Identifier in the Abstracts of Industry-Sponsored Manuscripts

LaVerne A. Mooney, Joseph F. Michalski, and Lorna Fay (United States)

MONDAY

Assessment of the Quality and Transparency of Research Reporting Endorsement by Brazilian Health Science Journals

Tais F. Galvao, Monica C. Roa, Leila Posenato Garcia, and Marcus T. Silva (Brazil)

MONDAY

Reporting Statistical Inference in Psychiatry, 1975-2015: An Analysis of Abstracts in Major Psychiatric Journals

Christopher Baethge, Markus Deckert, and Andreas Stang (Germany)

TUESDAY

A Scale for the Assessment of Non-systematic Review Articles (SANRA)

Christopher Baethge, Sandra Goldbeck-Wood, and Stephan Mertens (Germany, United Kingdom)

TUESDAY

Completeness of Reporting in Indian Qualitative Public Health Research: A Systematic Review of 20 Years of Literature

Myron A. Godinho, Nachiket Gudi, Maja Milkowska, Shruti Murthy, Ajay Bailey, and N.Sreekumaran Nair (India)

MONDAY

Reporting of Sex and Gender in Clinical Trial Protocols and Published Results

Thiyagu Rajakannan, Kevin Fain, Rebecca Williams, Tony Tse, and Deborah A. Zarin (United States)

TUESDAY

Adherence to Consolidated Standards of Reporting Trials (CONSORT) Guideline Items in Randomized Trials of Physical Activity Published in 5 Sports Medicine Journals

Daniel Umpierre, Lucas Helal, Patrícia Martins Bock, and Lucas Porto Santos (Brazil)

TUESDAY

Identification of Ethics Committees Based on Authors' Disclosures: Cross-sectional Study of Articles Published in the *European Journal of Anaesthesiology* and a Survey of Ethics Committees

Davide Zoccatelli, Martin R. Tramèr, and Nadia Elia (Switzerland)

Quality of the Literature

MONDAY

The Role of Supplementary Material in Journal Articles: Surveys of Authors, Reviewers, and Readers

Amy Price, Sara Schroter, Mike Clarke, and Helen McAneney (United Kingdom)

MONDAY

Editorial Evaluation, Peer Review, and Publication of Research Papers With and Without Online-Only Supplements: Quality vs Superior Tonnage

Annette Flanagan, Stacy Christiansen, Chris Borden, Demetrios N. Kyriacou, Caroline Sietmann, Elaine Williams, Larry Bryant, Jamie Reickel, and Allison Newton (United States)

TUESDAY

Readability of Open Access and Non-Open Access Articles Reporting Research Studies in Primary Health Care: A Cross-sectional Study

Shuhei Ichikawa, Kae Uetani, Yoshihito Goto, and Takanori Fujita (Japan)

Reporting Guidelines

MONDAY

Interventions to Improve Adherence to Reporting Guidelines: A Scoping Review

David Blanco de Tena-Dávila, Jamie Kirkham, Douglas G. Altman, David Moher, Isabelle Boutron, and Erik Cobo (Canada, France, Spain, United Kingdom)

TUESDAY

A Qualitative Assessment of the STROBE Extensions: Laying the Groundwork for Future Educational Interventions

Melissa K. Sharp and Darko Hren (Croatia)

MONDAY

Transparency and Completeness in the Reporting of Stakeholder Involvement in the Development and Reporting of Research Reporting Guidelines

Karen L. Woolley, Serina Stretton, and Lauri Arnstein (Australia, Japan, United Kingdom)

MONDAY

Evaluation of Reporting Guideline Implementation by Editors of Rehabilitation-Related Journals

Allen Heinemann, Leighton Chan, Helen Hoenig, Glenn Collins, and Jason Roberts (United States)

TUESDAY

Journal Support for ARRIVE Guidelines and Reporting Quality of Animal Welfare, Analgesia, or Anesthesia Articles

Vivian Leung, Frédéric Rousseau-Blass, and **Daniel S. J. Pang** (Canada)

Reproducible Research

TUESDAY

Association of Random Audits of Researchers With Improved Overall Quality of Research

Adrian G. Barnett, Nicholas Graves, and Pauline Zardo (Australia)

Research Methods

TUESDAY

Study Designs for the Evaluation of Biomarkers in Ovarian Cancer: A Systematic Review

Maria Olsen, Mona Ghannad, and Patrick M. Bossuyt (the Netherlands)

Statistics

MONDAY

Benefits and Barriers to Implementation of Statistical Review at a Veterinary Medical Journal: A Mixed-Methods Study

Alexandra Winter, Nicola DiGirolamo, and Michelle Giuffrida (Italy, United States)

MONDAY

Authors' Assessment of the Impact and Value of Statistical Review in a General Medical Journal: 5-Year Survey Results

Catharine Stack, Alicia Ludwig, A. Russell Localio, Anne Meibohm, Eliseo Guallar, John Wong, Deborah Cotton, Cynthia Mulrow, Jaya Rao, Mary Beth Schaeffer, Darren Taichman, and Christine Laine (United States)

Trial Registration

TUESDAY

Proportion of National Institutes of Health R01-Funded Clinical Trials Registered in ClinicalTrials.gov

Erick H. Turner, An-Wen Chan, Dan A. Oren, and Steven Bedrick (Canada, United States)

MONDAY

Clinical Trials and Tribulations: “The Registration Haze”

Denise M. Goodman, Karen E. Gutzman, and William F. Balistreri (United States)

TUESDAY

Adverse Event Reporting in Registered and Published Clinical Trials Focusing on Drug-Drug Interactions

Diana Jurić, Shelly Pranić, Ivančica Pavličević, and **Ana Marušić** (Croatia)

MONDAY

Adherence to the ICMJE Prospective Registration Policy Among Trials Published in High-Impact Specialty Society Journals

Anand D. Gopal, Joshua D. Wallach, Jenerius A. Aminawung, Gregg Gonsalves, Rafael Dal-Ré, Jennifer E. Miller, and Joseph S. Ross (Spain, United States)

Plenary Session Abstracts

Note: Abstracts reflect the status of the research at the time the abstracts were accepted for presentation.

Sunday, September 10, 2017 Common Problems in Peer Review and Scientific Publication

Bias Associated With Conflict of Interest and Peer Review

The Prevalence of Conflict of Interest Disclosures in Biomedical Research

Quinn Grundy,¹ Adam Dunn,² Florence Bourgeois,^{3,4} Lisa A. Bero¹

Objective Conflict of interest disclosures are used to indicate a risk of bias in biomedical research but studies examining their prevalence are out of date or focused on narrow clinical topics. Our aim is to estimate the prevalence of conflict of interest disclosures in biomedical research across disciplines and determine article characteristics associated with higher rates of disclosure.

Design We randomly sampled articles in Medline published from January 1, 2016, to December 31, 2016, in journals following the recommendations of the International Committee of Medical Journal Editors (ICMJE). There were no language restrictions. Non-peer-reviewed articles, including letters and news stories, were excluded. We developed a coding manual to classify the reported conflicts of interest and sources of study funding based on the National Academies of Medicine definition of conflict of interest and the ICMJE disclosure form. Independently, 2 researchers piloted the coding manual on a random sample, resolving discrepancies through verification and discussion.

Results After sampling 1650 articles, 1002 articles met our inclusion criteria. We found that 22.9% (95% CI, 20.3%-25.6%) disclosed a conflict of interest, 63.6% (95% CI, 60.5%-66.6%) disclosed no conflicts, and 13.6% (95% CI, 11.5%-15.7%) did not include a disclosure statement (**Table 1**). Articles focused on drugs, devices, or surgical procedures were significantly more likely to include authors with reported conflicts of interests (71 of 267 [26.6%]) than other empirical articles (64 of 415 [15.4%]) (difference, 11.2%; 95% CI, 3.3%-19.1%). Disclosure statements were inconsistent: we noted 130 different ways of stating there were no conflicts of interest, ranging from “None declared” to “Nothing to declare” to “No relevant conflicts” to statements

63 words long. Furthermore, 90 of 228 articles (39.4%) with statements contained extraneous biographical information not addressing conflicts of interest.

Conclusions Many articles published in journals following the ICMJE recommendations fail to include disclosure statements. Just more than 1 in 5 biomedical articles report a relevant conflict of interest, which is generally consistent with a 2003 review that found that 23% to 28% of academic investigators receive funding from industry, suggesting this may be an underestimate. In current practice, conflict of interest statements are unstructured and inconsistently reported, precluding automatic extraction and analysis of conflict of interest statements.

¹Charles Perkins Centre, Faculty of Pharmacy, The University of Sydney, Sydney, New South Wales, Australia; quinn.grundy@sydney.edu.au; ²Australian Institute of Health Innovation, Macquarie University, Sydney, New South Wales, Australia; ³Harvard Medical School, Harvard University, Cambridge, MA, USA; ⁴Division of Emergency Medicine, Boston Children's Hospital, Boston, MA, USA

**Table 1. Prevalence of Author Conflict of Interest (COI)
Disclosures**

Prevalence	Author COI Proportion, No. (%) (95% CI by Clopper-Pearson Exact)		
	Yes	No	Missing
Empirical articles	135 of 682 (19.8%) (16.9-23.0)	462 of 682 (67.7%) (64.1-71.2)	85 of 682 (12.5%) (10.1-15.2)
Drug-focused	39 of 124 (31.5%) (23.4-40.4)	69 of 124 (55.6%) (46.5-64.6)	16 of 124 (12.9%) (7.6-20.1)
Device-focused	27 of 121 (22.3%) (15.2-30.8)	75 of 121 (62.0%) (52.7-70.7)	19 of 121 (15.7%) (9.7-23.4)
Both drug and device	5 of 22 (22.7%) (7.8-45.4)	16 of 22 (72.7%) (49.8-89.3)	1 of 22 (4.5%) (0.1-22.8)
Neither drug nor device	64 of 415 (15.4%) (12.1-19.3)	302 of 415 (72.8%) (68.3-76.8)	49 of 415 (11.8%) (8.9-15.3)
Commentaries, editorials, and narrative reviews	91 of 290 (31.4%) (26.1-37.1)	150 of 290 (51.7%) (45.8-57.6)	49 of 290 (16.9%) (12.8-21.7)
Systematic reviews and meta-analyses	3 of 30 (10.0%) (2.1-26.5)	25 of 30 (83.3%) (65.3-94.4)	2 of 30 (6.7%) (0.8-22.1)
All articles	229 of 1002 (22.9%) (20.3-25.6)	637 of 1002 (63.6%) (60.5-66.6)	136 of 1002 (13.6%) (11.5-15.9)

Conflict of Interest Disclosures: Dr Bero is a Peer Review Congress Advisory Board Member but was not involved in the review or decision for this abstract.

The Influence of Industry Funding and Other Financial Conflicts of Interest on the Outcomes and Quality of Systematic Reviews

Camilla Hansen,^{1,2} Andreas Lundh,^{1,3} Kristine Rasmussen,² Tove Faber Frandsen,⁴ Peter C. Gøtzsche,² Asbjørn Hróbjartsson¹

Objective Funding of systematic reviews by drug and device companies and other financial conflicts of interest among authors may have an impact on how the reviews are conducted. The aim of this study was to investigate if financial conflicts of interest are associated with results, conclusions, and methodological quality of systematic reviews.

Design This is a Cochrane methodology review. We searched PubMed, EMBASE, and the Cochrane Methodology Register as well as the reference lists of included studies and Web of Science for studies citing the included studies. We included observational studies of any design that investigated samples of systematic reviews with and without industry funding or other financial conflicts of interest, published up to November 2016. For studies to be eligible, they had to investigate at least 1 of our outcomes: effect size estimates, statistically favorable results, favorable conclusions, and methodological quality. Two review authors independently extracted data and assessed risk of bias in relation to study inclusion, data extraction, and comparability of the investigated systematic reviews. We reported our findings on effect size estimates qualitatively. We calculated pooled risk ratios (RRs) with 95% confidence intervals for statistically favorable results, favorable conclusions, and methodological quality.

Results Nine observational studies with a total of 983 systematic reviews of drug studies and 15 systematic reviews of device studies were included. Effect size estimates and frequency of statistically favorable results were similar between systematic reviews with and without financial conflicts of interest (**Table 2**). Systematic reviews with financial conflicts of interest more often had favorable conclusions compared with systematic reviews without financial conflicts of interest (RR, 1.96; 95% CI, 1.23-3.13).

Conclusions Systematic reviews with financial conflicts of interest related to drug and device companies more often have favorable conclusions and to some degree lower methodological quality compared with systematic reviews without financial conflicts of interest. It remains unclear whether financial conflicts of interest have an impact on the results.

¹Center for Evidence-Based Medicine, Odense University Hospital and University of Southern Denmark, Odense, Denmark, camilla.hansen3@rsyd.dk; ²Nordic Cochrane Centre, Rigshospitalet, Copenhagen, Denmark; ³Department of Infectious Diseases, Hvidovre Hospital, Hvidovre, Denmark; ⁴Department of Design and Communication, University of Southern Denmark, Odense, Denmark

Table 2. Systematic Reviews With Financial Conflicts of Interest Compared With Systematic Reviews Without

Outcome	Comparative Risk		Effect Estimate
	Industry n/N (%)	Nonindustry n/N (%)	
Results			
Estimated effect sizes ^a	NA	NA	z Score: 0.46 (<i>P</i> = .64)
Statistically favorable results	27 of 49 (55)	49 of 75 (65)	RR, 0.84 (95% CI, 0.62-1.14)
Favorable conclusions	163 of 200 (82)	93 of 199 (47)	RR, 1.96 (95% CI, 1.23-3.13)
Methodological quality ^b			
Appropriate search methods	94 of 145 (65)	124 of 157 (79)	RR, 0.72 (95% CI, 0.49-1.06)
Appropriately selected studies	69 of 145 (48)	98 of 157 (62)	RR, 0.68 (95% CI, 0.44-1.06)
Appropriately combined studies	75 of 145 (52)	92 of 157 (59)	RR, 0.90 (95% CI, 0.70-1.14)
Had conclusions supported by the data	40 of 81 (49)	65 of 100 (65)	RR, 0.86 (95% CI, 0.62-1.21)
Assessed risk of bias	56 of 145 (39)	104 of 157 (66)	RR, 0.47 (95% CI, 0.23-0.95)
Interpreted results in light of risk of bias	49 of 127 (39)	69 of 120 (58)	RR, 0.68 (95% CI, 0.53-0.87)

Abbreviations: NA, not applicable; RR, risk ratio.
^aMeasured as pooled z score. A z score expresses the number of standard deviations a value differs from the mean.
^bRR < 1 indicates that systematic reviews with financial conflicts of interest have lower methodological quality.

Conflict of Interest Disclosures: Dr Gøtzsche is a Peer Review Congress Advisory Board member but was not involved in the review or decision for this abstract.

Analysis of Uptake and Outcome in Author-Selected Single-blind vs Double-blind Peer Review at Nature Journals

Elisa De Ranieri¹; Barbara McGillivray²; Sowmya Swaminathan³; Michelle Samarasinghe⁴; Leah Gruen⁴

Objective Double-blind peer review might avoid referee bias. The aims of this study were to analyze the demographics of corresponding authors choosing double-blind peer review and to identify differences in editorial outcome depending on review model.

Design Data include direct submissions and transfers received between March 2015 and February 2017 by 25 Nature-branded journals. The authors chose either single- or double-blind review, and the editors were aware of the choice before taking any decisions. We analyzed direct submissions to study the uptake of double-blind review in relation to gender, country, and institutional prestige of the corresponding author. We analyzed all submissions to study the editorial outcome in relation to review model. The gender (male, female, or not available) of the corresponding authors was determined from their first name using a third-party service (Gender API), discarding results with less than 80%

confidence. The prestige of corresponding author's institutions was measured by normalizing the institution's name using the Global Research Identifier Database (GRID) and dividing institutions in 3 prestige groups using the 2016 Times Higher Education (THE) ranking. We used descriptive statistics for data exploration; we tested our hypotheses using Pearson's χ^2 and binomial tests.

Table 3. Data Concerning the 2 Review Models Based on Several Attributes of the Submission or of the Manuscript's Corresponding Author^a

Attribute	Double Blind, No. (%)	Single Blind, No. (%)	P Value
Direct submissions ^b	12,631 (12)	93,742 (88)	NA
Nature ^b	2782 (14)	17,624 (86)	<2.2e-16
Sister journals ^b	8053 (12)	57,181 (88)	
Nature Communications ^b	3900 (9)	38,914 (91)	
Gender of corresponding author ^c			
Female	1506 (10)	12,943 (90)	.62
Male	7271 (11)	61,536 (89)	
Institution group ^d			
1	240 (4)	5818 (96)	<2.2e-16
2	1663 (8)	19,295 (92)	
3	4174 (13)	27,730 (87)	
Country ^e			
Australia	274 (10)	2366 (90)	<2.2e-16
Canada	259 (9)	2581 (91)	
China	3626 (22)	13,148 (78)	
France	278 (8)	3334 (92)	
Germany	350 (5)	6079 (95)	
India	711 (32)	1483 (68)	
Japan	933 (15)	5248 (85)	
South Korea	643 (12)	3089 (88)	
United Kingdom	509 (7)	6656 (93)	
United States	2298 (7)	30,184 (93)	
Other	2750 (12)	19,574 (88)	
Out to review decision ^f			
Sent	1242 (8)	25,985 (23)	<2.2e-16
Not sent	13,493 (92)	87,734 (77)	
Decision after review ^g			
Accepted	242 (25)	8692 (44)	<2.2e-16
Rejected	732 (75)	11,040 (56)	

^aWhen applicable, we show *P* values from hypothesis tests performed to test the null hypothesis that there is no association between review model and each attribute (eg, journal category).

^bData set for overall uptake: 106,373 direct submissions. Sister journals are Nature-branded journals, excluding *Nature* and *Nature Communications*.

^cData set for gender analysis: 83,256 direct submissions.

^dInstitution groups are defined to include institutions with a Times Higher Education (THE) rank between 1 and 10 (group 1), 11 and 100 (group 2), and above 101 (group 3). Submissions from institutions without a THE ranking are not included. Data set for institutional prestige analysis: 58,920 direct submissions.

^eThe itemized countries are responsible for 80% (85,098) of direct submissions. All other countries are grouped under "Other."

^fData set for out-to-review statistics: 128,454 direct submissions and transfers.

^gData set for final outcome statistics: 20,706 direct submissions and transfers.

Results Out of 128,454 papers, 106,373 were direct submissions, of which 12% were submitted double-blind review (**Table 3**). We found a small but significant association between journal tier and review type. We had gender information for 50,533 corresponding authors (in 83,256 submissions) and found no statistically significant difference in the distribution of peer-review model between males and females. We had 58,920 records with normalized institutions and a THE rank, and we found that corresponding authors from the less prestigious institutions are more likely to choose double-blind review. In the 10 countries with the highest number of submissions, we found a small but significant association between country and review type. China and the United States had a preference for double- and single-blind review, respectively. The outcome at both first decision and postreview was significantly more negative (ie, a higher likelihood for rejection) for double-blind than single-blind reviewed papers, and we attribute this to differences in the quality of the studies.

Conclusions Authors choose double-blind review more frequently when they submit to more prestigious journals, when they are affiliated with less prestigious institutions, or when they are from specific countries. The double-blind option is also linked to less successful editorial outcomes.

¹Springer Nature, London, UK, e.deranieri@nature.com; ²Alan Turing Institute and Department of Theoretical and Applied Linguistics, Faculty of Modern and Medieval Languages, University of Cambridge, UK; Springer Nature, San Francisco, CA, USA; ³Springer Nature, New York, NY, USA

Conflict of Interest Disclosures: All authors are or have been employed by Springer Nature, which owns and publishes the Nature-branded journals.

Gender and Age Bias in Peer Review in Earth and Space Science Journals

Jory Lerback,¹ Brooks Hanson²

Objective The American Geophysical Union (AGU) publishes, 20 journals with approximately 6000 articles and 24,000 reviews annually. We studied the gender differences and dynamics in publishing and reviewing. This has been studied in other disciplines, but these studies have mostly assigned gender to first names (we have self-reported gender), had smaller sample sizes, and/or have not accounted for age.

Design We analyzed membership demographic data and editorial data from the AGU from 2012 to 2016. We analyzed activities in the publications database, looking at demographic data for 23,985 distinct reviewers, 29,927 first authors, 97,120 reviewer suggestions by authors, and 151,484 reviewer invitations by editors. Age is important to include because the proportion of women researchers decreases as age increases; accounting for age is needed to reveal some otherwise hidden gender differences.

Results Female first authors had higher acceptance rates than men across all age cohorts (61 vs 58%; χ^2_1 [n = 29,187] =

Table 4. Proportion of Female Individuals in Peer Review Interactions

Actors	Population Analyzed	% Female of Population (All Ages)	Females in Population Analyzed, No.	Males in Population Analyzed, No.
	Distinct members	29.6	22,570	55,098
	Distinct published first authors	26.9	2661	7248
	Distinct published authors	23.6	5433	17,622
Female first authors	Coauthors	21.7	4069	14,641
Female corresponding authors	Reviewer suggestions	21.5	4616	16,832
Female editor	Reviewer invitations	22.1	5037	17,772
Male first authors	Coauthors	16.0	8811	46,291
Male corresponding authors	Reviewer suggestions	15.8	11,991	63,681
Male editors	Reviewer invitations	17.5	22,474	105,782

20.057). Women make up 27% of first authors ($n = 9,909$), 24% of all authors ($n = 18,710$), and 30% of AGU membership ($n = 77,668$) (**Table 4**). Despite this, women were not utilized as reviewers (21%) as much as expected based on these rates (χ^2_1 first authors = 145.396 [$n = 33,395$]; $P < .001$; χ^2_1 all authors = 50.958 [$n = 47,081$]; $P < .001$; χ^2_1 members = 629.231 [$n = 101,694$]; $P < .001$, respectively). Although the proportion of female reviewers increased from 2012 to 2016, this gap persisted and was consistent throughout age cohorts of the suggested reviewers. This difference began with authors, who suggested male reviewers more than expected (male authors suggested 16% female reviewers [$n=75,672$]; female authors suggested 22% ($n=21,488$). Male editors subsequently invited only 18% female reviewers, whereas female editors invited 22%. This difference in suggestions partly parallels coauthor networks, in which male first authors tend to have other males as collaborators (16% [$n=55,102$]), whereas female first authors had collaborators that more closely represented the gender-age distribution of the research population (22% [$n=18,710$]).

Conclusions We found that women are not being included in activities related to peer review processes as frequently as their male peers in Earth and space journals.

¹Department of Geology and Geophysics, University of Utah, Salt Lake City, UT, USA, jory.lerback@gmail.com; ²American Geophysical Union, Washington, DC, USA

Conflict of Interest Disclosures: None reported.

Additional Contributions: We thank the American Geophysical Union for providing data.

Bias in Reporting and Publication of Research

Augmenting Systematic Reviews With Information From ClinicalTrials.gov to Increase Transparency and Reduce Bias

Stacey Springs,¹ Gaelen Adam,¹ Thomas Trikalinos,¹ John W. Williams Jr.,² Jennifer Eaton,² Megan Von Isenburg,² Jennifer M. Gierisch,² Lisa M. Wilson,³ Ritu Sharma,³ Sydney M. Dy,³ Julie M. Waldfoegel,³ Karen A. Robinson,³ Meera Viswanathan,⁴ Jennifer Cook Middleton,⁴ Valerie L. Forman-Hoffman,⁴ Elise Berliner,⁶ Robert M. Kaplan^{5,6}

Objective Prospective registration of clinical trials may improve transparency and reduce bias associated with selective reporting. Our objective was to evaluate the impact of access to and integration of information from ClinicalTrials.gov on the conclusions of systematic reviews in 5 clinical areas.

Design Teams of systematic reviewers searched ClinicalTrials.gov for studies relevant to 5 different ongoing systematic reviews: Effectiveness of Treatment Options for the Prevention of Complications and Treatment of Symptoms of Diabetic Peripheral Neuropathy (DPN), Management of Infertility, Omega-3 Fatty Acids and Cardiovascular Disease, Strategies to Improve Mental Health Care for Children and Adolescents, and Tympanostomy Tubes in Children with Otitis Media. A semi-automated approach to matching studies using EndNote was not feasible owing to lack of standardization of format and location of the registry identification number in published reports. Teams compared trials, and information on trials, found from searches of other sources and determined whether information uniquely found in ClinicalTrials.gov changed confidence in evidence and review conclusion.

Results Across all topics, 24% (101 of 419) of all included trials were registered in ClinicalTrials.gov; 38% (95 of 251 total registry records found) did not have results published in peer reviewed literature; and of trials with published and registry reported results, 63% (124 of 198) of outcomes matched in the publication and ClinicalTrials.gov records (**Table 5**) Despite the additional trials found in the searches of ClinicalTrials.gov, the strength of evidence and conclusions in each systematic review were unchanged, primarily owing to missing results of most of the additional trials found.

Conclusions Across topic areas, only 24% (101 of 419) were registered in ClinicalTrials.gov and 38% (95 of 251) of studies did not have results published in peer-reviewed literature. The potential impact of this missing information on the conclusions of systematic reviews is unknown. When there were both ClinicalTrials.gov records and publications, 37% (74 of 198) of outcome measures did not match, raising a concern about bias owing to selective outcome reporting. It appears that prespecification of a primary outcome variable in ClinicalTrials.gov does not inhibit reporting other outcomes in publications. New rules requiring outcome

Table 5. Summary of ClinicalTrials.gov Registration for Published and Unpublished Results in 5 Topic Areas

Review Topic (n trials)	CT.gov Record For Published Trial		CT.gov Registered Trials Without Public Results Reporting		Outcome Measures Reported, No.		
	Yes	No	Results Not Published/Total Studies Found in Registry (y since completion)	Results Not Registered in CT.gov	CT.gov and Publication	CT.gov Only	Publication Only
Effectiveness of treatment options for diabetic peripheral neuropathy (n= 106)	53	53	23/53 ≤ 3 y: 10 >3 y: 13	36	18 outcomes (n=12 studies)	5 outcomes Results: 1 (n=5 studies)	22 (n=16 studies)
Management of infertility (n=24)	12	12	4/94 ≤ 3 y: 2 >3 y: 2	11	11 outcomes	3 outcomes Results: 0 (n= 3 studies)	6 outcomes (n=4 studies)
Omega-3 fatty acids and cardiovascular disease (n=98)	26	72	43/69 ≤ 3 y: 34 >3 y: 9	43	71 outcomes (n=26 studies)	2 outcomes Results: 0 (n=2 studies)	25 outcomes (n=12 studies)
Mental health care for children and adolescents (n=13)	4	9	3/6 All ongoing	NA	Not evaluated	0	13 outcomes (n=12 studies)
Tympanostomy tubes in children with otitis media (n=178)	6	172	22/28 ≤ 3 years: 17 >3 years: 5	20	24 outcomes (n=6 studies)	8 outcomes Results: 0 (n=2 studies)	3 outcomes (n=2 studies)

measure specification and reporting should be considered. Journals and indexing tools could facilitate the inclusion of information from ClinicalTrials.gov into systematic review by adopting a more standardized format for listing the ClinicalTrials.gov identification number.

¹Brown University Evidence Based Practice Center, Providence, RI, USA; ²Duke University Evidence Based Practice Center, Durham, NC, USA; ³Johns Hopkins University Evidence Based Practice Center, Baltimore, MD, USA; ⁴RTI/UNC Evidence Based Practice Center, Raleigh, NC, USA; ⁵Agency for Healthcare Research and Quality, Rockville, MD, USA, elise.berliner@ahrq.hhs.gov; ⁶Stanford University, Stanford, CA, USA

Conflict of Interest Disclosures: None reported.

Funding/Support: This project was funded by the Agency for Healthcare Research and Quality.

Role of the Funder/Sponsor: Representatives from the Agency for Healthcare Research and Quality (AHRQ) served as a Contracting Officer's Technical Representative and provide technical assistance during the conduct of the project and provided comments on draft versions of the project. AHRQ did not directly participate in the literature search, determination of study eligibility criteria, data analysis or interpretation.

Disclaimer: The findings and conclusions in this abstract are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ and no statement in this abstract should be construed as an official position of AHRQ or the US Department of Health and Human Service.

Bias Associated With Publication of Interim Results of Randomized Trials: A Systematic Review

Steven Woloshin,¹ Lisa M. Schwartz,¹ Pamela J. Bagley,² Heather B. Blunt,² Brian White³

Objective Publication of interim results from ongoing randomized clinical trials may generate substantial interest because they are new and often promising. Final results,

however, may not confirm early promise and may receive less attention. Our objective was to describe the publication of interim results from randomized clinical trials and to compare the prominence and consistency with final publications.

Design We conducted a PubMed search (2006-2015) for interim publications of randomized clinical trials, including the terms *interim*, *not mature*, or *immature* in the title or abstract. We used registration numbers and author names to search PubMed, ClinicalTrials.gov, and Web of Science through 2016 for final publications (authors were contacted if none identified) and determined each publication's journal Impact Factor and Altmetric rating (ie, news and social media attention). Two researchers confirmed interim and final publication pairing and abstracted data.

Results Of 1267 publications screened, 613 reported interim results (excluding completed pilot studies, protocols, and cancer trials reporting an interim result for the secondary outcome [overall survival] but with a final primary outcome result [progression-free survival]). Seventy-two percent (442 of 613) of these publications reported on trials stopped early (for benefit [105], harm [67], futility [224], other problems [46]). The remaining 171 ongoing trials reported interim efficacy or safety results. Forty percent (68 of 171) stated the reason for interim publication was a protocol-specified preplanned analysis; a few (6% [10 of 171]) stated other reasons (eg, response to release of results about the same intervention), but most (54% [93 of 171]) stated none. The 171 interim publications were mostly in oncology (28% [48]), surgery (18% [30]), or cardiology (11% [18]); 59% (101) had active controls, and 13% (23) tested noninferiority. The most commonly stated funding sources were solely industry (36% [61]), partly industry (10% [17]), government (18% [30]), and foundation or university (17% [29]). Final results were published for 57% (90) of the 158 trials where sufficient time elapsed for final publication (eg, ≥1 year beyond registry-specified study completion date). Most abstract conclusions

(85% [61]) did not change qualitatively for the 72 pairs of interim and final publications reporting the same primary outcome results, while 15% changed: 8% (6) became weaker (eg, changed from “superior” to “not superior”), and 7% (5) became stronger. Interim and final publications had similar prominence in terms of Impact Factor and Altmetric rating (**Table 6**).

Conclusions Frequent nonpublication of final results may cause bias because true treatment effects often remain unknown. Final publications, when available, have as much journal and media prominence as interim publications but may reach qualitatively different conclusions. Journals should publish fewer interim results (especially when not prespecified) and commit to making the final results known when they do.

[†]The Center for Medicine in the Media, Dartmouth Institute for Health Policy and Clinical Practice, Lebanon, NH, USA, lisa.schwartz@dartmouth.edu; [‡]Biomedical Libraries, Dartmouth College, Hanover, NH, USA; [§]Atlanticare Regional Medical Center, Pomona, NJ, USA

Conflict of Interest Disclosures: Drs Schwartz and Woloshin have served as medical experts in testosterone litigation. No other conflicts were reported.

Table 6. Key Performance Indicators for the 72 Pairs of Interim and Final Publications Both Reporting Primary Outcome Results

Performance Indicator	Interim	Final	P value ^a
High impact factor journal (>20)	26% (19 of 72)	29% (21 of 72)	.73
Top 5 impact factor general journals	19% (14 of 72) ^b	14% (10 of 72)	.40
Both in Top 5	NA	35% (5 of 14)	NA
Altmetric-rating median [IQR] ^c	7 [2-38]	6 [2-42]	.86

Abbreviations: IQR, interquartile range; NA, not applicable.
^aP values for paired data (McNemar test for dichotomous and sign test for continuous variables).
^bAll published in the *New England Journal of Medicine* or *Lancet*.
^cAltmetric ratings were only available for 38 interim and 53 final publications.

Identification and Classification of Spin in Clinical Studies Evaluating Biomarkers in Ovarian Cancer: A Systematic Review

Mona Ghannad,^{1,2} Maria Olsen,^{1,2} Patrick M. Bossuyt¹

Objective The objective of this systematic review was to document and classify spin or overinterpretation, as well as facilitators of spin, in recent clinical studies evaluating performance of biomarkers in ovarian cancer.

Design We searched PubMed systematically for all studies published in 2015. Studies eligible for inclusion described 1 or more trial designs for identification and/or validation of prognostic, predictive, or diagnostic biomarkers in ovarian cancer. Reviews, animal studies, and cell line studies were excluded. All studies were screened by 2 reviewers. To document and characterize spin, we collected information on the quality of evidence supporting the study conclusions, linking the performance of the marker to outcomes claimed.

Results In total, 1026 potentially eligible articles were retrieved by our search strategy, and 345 studies met all eligibility criteria and were included. The first 200 studies, when ranked according to publication date, will be included in our final analysis. Data extraction was done by one researcher and validated by a second. Specific information extracted and analyzed on study and journal characteristics, key information on the relevant evidence in methods, and reporting of conclusions claimed for the first 50 studies is provided here. Actual forms of spin and facilitators of spin were identified in studies trying to establish the performance of the discovered biomarker. Actual forms of spin identified as shown (**Table 7**) were: (1) other purposes of biomarker claimed not investigated (18 of 50 studies [36%]); (2) incorrect presentation of results (15 of 50 studies [30%]); (3) mismatch between the biomarker’s intended clinical application and population recruited (11 of 50 studies [22%]); (4) mismatch between intended aim and conclusion (7 of 50 studies [14%]); and (5) mismatch between abstract conclusion and results presented in the main text (6 of 50 studies [12%]). Frequently observed facilitators of spin were: (1) not clearly prespecifying a formal test of hypothesis (50 of 50 studies [100%]); (2) not stating sample size calculations (50 of 50 studies [100%]); (3) not prespecifying a positivity threshold of continuous biomarker (17 of 43 studies [40%]); (4) not reporting imprecision or statistical test for data shown (ie, confidence intervals, P values) (12 of 50 studies [24%]); and (5) selective reporting of significant findings between results for primary outcome reported in abstract and results reported in main text (9 of 50 studies [18%]).

Conclusions Spin was frequently documented in abstracts, results, and conclusions of clinical studies evaluating performance of biomarkers in ovarian cancer. Inflated and selective reporting of biomarker performance may account for a considerable amount of waste in the biomarker discovery process. Strategies to curb exaggerated reporting are needed to improve the quality and credibility of published biomarker studies.

¹Academic Medical Center Amsterdam, Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam Public Health Research Institute, Amsterdam, the Netherlands, m.ghannad@amc.uva.nl; ²Centre de Recherche Épidémiologie et Statistique Sorbonne Paris Cité (CRESS-UMR1153), Université Paris Descartes, Centre d’épidémiologie Clinique, Hôpital Hôtel-Dieu, Paris, France

Conflict of Interest Disclosures: None reported.

Funding/Support: This project has received funding from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 676207.

Role of the Funder/Sponsor: The funder/sponsor had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the abstract.

Table 7. Prevalence of Actual Forms of Spin in Clinical Studies Evaluating Performance of Biomarkers in Ovarian Cancer

Actual Forms of Spin	All Studies (n=50)	Examples of Actual Forms of Spin
Other purposes of biomarker claimed not investigated	18 (36%)	Example 1: Potential use of BM for screening not the aim or investigated in the study. <i>"With a short analysis time and inclusion of novel markers for early ovarian cancer detection, this platform shows strong promise as a potential point of care screening method for ovarian cancer, where patients could receive results promptly enough to be referred to transvaginal sonography in the same visit. Reduced costs and easier accessibility to results could also assist in longitudinally monitoring biomarker values over time, which has shown some promise in helping detect early stage ovarian cancer."</i>
Incorrect presentation of results	15 (30%)	Example 1: Alternative facts: describe negative association and odds ratio of 0.532 as "protective role." <i>"In addition to this finding, we observed that rs3814113 on 9p22 may play a protective role from the development of serous histological subtypes of ovarian carcinoma."</i> Example 2: Claim effect despite statistically insignificant results. Example 3: Claim effect despite not providing imprecision or statistical test (confidence intervals or <i>P</i> values) between different biomarker models tested or patient groups (subgroups).
Mismatch between biomarker's intended clinical application and the population recruited	11 (22%)	Example 1: The use of healthy controls for the performance of the BM evaluating diagnostic, prognostic, or predictive treatment response. Similarly recruitment of symptomatic women for the performance of the BM in screening or risk.
Mismatch between intended aim and conclusion	7 (14%)	Example 1: Extrapolation of preclinical study results to clinical application. Example 2: Use of causal language for BM(s): being assessed despite the use of a nonrandomized design.
Mismatch between abstract conclusion and results presented in the main text	5 (10%)	Example 1: Leap from association to genetic risk factor. <i>"Despite the relatively small sample size of cases and controls, our studies confirmed some of the previously demonstrated GWAS single-nucleotide polymorphisms as genetic risk factors for epithelial ovarian tumors."</i>
Mismatch between results reported in abstract and results reported in main text	3 (6%)	Example 1: The direction of the association in the results is negative, which they interpret as "protective." However, the abstract indicates "significant" association, implying positive. Example 2: The reported HRs reported in the abstract does not match the HRs reported in the main text. In main text the HRs for PFS and OS are reported for each quartile. Unclear if what is reported in abstract is the overall HRs.
Other benefits claimed that is not prespecified and/or investigated	2 (4%)	Example 1: Reduced costs and easier accessibility to results.

Abbreviations: BM, biomarker; HR, hazard ratio; OS, overall survival; PFS, progression-free survival.

Spin in Published Biomedical Literature: A Systematic Review

Quinn Grundy,¹ Kellia Chiu,¹ Lisa A. Bero¹

Objective To explore the nature and prevalence of spin in the biomedical literature.

Design In a systematic review and meta-analysis, we searched MEDLINE, PreMEDLINE, Embase, Scopus, and handsearched reference lists for all articles published between 1946 and 24 November 2016 that included the quantitative measurement of spin in the biomedical literature for at least 1 outcome. Two independent coders extracted data on the characteristics of articles and included studies, methods for assessing spin, and all spin-related results. The data were heterogeneous; results were grouped inductively into outcome-related categories. We had sufficient data to use meta-analysis to analyze the association of industry sponsorship of research with the presence of spin.

Results We identified 4219 articles after removing duplicates and included 35 articles that investigated spin: clinical trials (23/35, 66%); observational studies (7/35, 20%); diagnostic accuracy studies (2/35, 6%); and systematic reviews and meta-analyses (4/35, 11%), with some articles including multiple study designs. The nature and manifestations of spin varied according to study design. We grouped results into the following categories: prevalence of spin, level of spin, factors associated with spin, and effects of spin on readers' interpretations. The highest, but also greatest variability in the prevalence of spin was present in trials (median, 57% of main texts containing spin; range, 19%-100% across 16

articles). Source of funding was hypothesized to be a factor associated with spin; however, the meta-analysis found no significant association, possibly owing to the heterogeneity of the 7 included articles.

Conclusions Spin appears to be common in the biomedical literature, though this varies by study design, with the highest rates found in clinical trials. Spin manifests in diverse ways, which challenged investigators attempting to systematically identify and document instances of spin. Widening the investigation of factors contributing to spin from characteristics of individual authors or studies to the cultures and structures of research that may incentivize or deincestivize spin, would be instructive in developing strategies to mitigate its occurrence. Further research is also needed to assess the impact of spin on readers' decision making. Editors and peer reviewers should be familiar with the prevalence and manifestations of spin in their area of research to ensure accurate interpretation and dissemination of research.

¹Charles Perkins Centre, Faculty of Pharmacy, The University of Sydney, Sydney, New South Wales, Australia, quinn.grundy@sydney.edu.au

Conflict of Interest Disclosures: Dr Bero is a member of the Peer Review Congress Advisory Board but was not involved in the review or decision for this abstract.

Integrity and Misconduct

Summary Effect Sizes in Meta-analyses After Removal of Retracted Studies From the Pool of Primary Studies

Daniele Fanelli,¹ David Moher^{2,3}

Objective This study aimed to assess the magnitude and direction of change of summary effect size in meta-analyses (MAs) after retracted papers are removed from the pool of primary studies.

Design We aimed to identify a homogeneous sample of recent MAs that contained, among primary studies, 1 or more studies that were later retracted, to compare pooled odds ratios with and without such studies. On December 16, 2016, we retrieved all retracted publications recorded in the Web of Science (WOS) and then retrieved a list of records that cited these retracted publications. We selected all records containing “meta-analysis” or “systematic review” in the title, abstract, or keywords and then restricted the initial list of potentially relevant titles to records published in 2016. The full text of these studies was retrieved and inspected for selection based on the following exclusion criteria: limited to a systematic review and not a formal MA; not a standard MA (ie, a weighted pooled summary of primary studies, which excludes network MAs, genome-wide association studies, and MAs of functional magnetic resonance imaging, microarray, and genomic data); does not contain primary summary data in the full text (ie, the retracted cited article is not among primary studies of the MA); or does not use odds ratio–convertible metrics (including risk difference, proportion, mean, or other unusual metrics designed for the specific purposes of a study).

Results A total of 3834 records of potentially retracted articles were identified in WOS. We retrieved 83,946 records that cited these potentially retracted publications; from these, we identified 1433 records containing “meta-analysis” or “systematic review” in the title, abstract, or keywords. Of the 109 potentially relevant MAs published in 2016, 17 did not match any exclusion criteria and were included in this study. Each of these MAs had included in its weighted summary 1 retracted study. Three pairs of MAs cited the same retracted study; therefore, the number of distinct retraction events covered in our sample is 14. All MAs had been authored by independent research teams, and only 1 author appeared in 2 MAs. Two MAs were published in the Cochrane Database of Systematic Reviews, and 15 were published in different journals that were classified by WOS in different biomedical fields, from molecular biology to surgery. Additional analyses are ongoing.

Conclusions The 17 MAs included in the study are representative of multiple biomedical research areas and retraction events. For each of these MAs, we will calculate summary effect size with and without the retracted primary study and obtain a ratio of odds ratios across the sample.

Pooled results will yield a preliminary estimate of the possible impact that retractions may have on the biomedical literature.

¹Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Palo Alto, CA, USA, email@danielefanelli.com; ²Centre for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, The Ottawa Hospital, Ottawa, Ontario, Canada; ³School of Epidemiology, Public Health and Preventive Medicine, University of Ottawa, Ottawa, Canada

Conflict of Interest Disclosures: Dr Moher is a member of the Peer Review Congress Advisory Board, but was not involved in the review or decision for this abstract. No other disclosures were reported.

Acknowledgments: Research assistant Julie Wong helped collect the raw data.

Assessing the Outcomes of Introducing a Digital Image Quality Control Review Into the Publication Process for Research Articles in Physiology Journals

Rita Scheman,¹ Christina N. Bennett¹

Objective To address concerns that American Physiological Society (APS) journals were publishing photographs, mostly Western blots and DNA/RNA gel images, that had inappropriate and/or undeclared modifications, we introduced a postacceptance but prior to early-view publication quality control (QC) procedure in 7 journals using digital image forensic tools to check for image splicing, duplication, extreme contrast, and selective editing. We sought to assess whether the QC check effectively identified and corrected images with modifications prior to publication, and whether corresponding authors we have queried about their images after a QC check have submitted another manuscript without generating another QC query.

Design We assessed the number of QC cases queried per year, image modifications identified, and manuscript outcomes categorized as no revision, revision, corrigendum, rejection, or retraction between 2009 and 2016. We also assessed the number of subsequent submissions of unique manuscripts by corresponding authors involved in initial QC queries and the outcomes of those subsequent QC checks. We report results for 3 time periods: 2009, when no article underwent the QC check unless concerns arose during figure preparation for final publication; 2010 to 2012, when the procedure was introduced in 7 APS journals one at a time; and 2013 to 2016 when all 7 journals used the QC check process.

Results The QC checks were performed on 1.1% of research articles in 2009, 5.9% in 2010 to 2012, and 6.5% in 2013 to 2016. Implementation of the QC check reduced the number of corrigenda published (from 22/25 queries in 2009 to 0/71 in 2016) with a reciprocal increase in the number of revisions prior to publication (3/25 in 2009 to 65/71 in 2016). Since 2009, only 23 of 733 articles contained image modifications serious enough to rescind acceptance or retract the early view version. Since 2013, the proportion of QC queries has decreased 0.7% (95% CI 1.2%–0.3%) each year ($P = .03$).

Fifty-eight percent (190/326) of corresponding authors who received QC queries from us between 2013 and 2015 submitted another manuscript for publication to one of the journals, and only 8 were involved in a subsequent QC query.

Conclusions Implementing a QC check for review of image modifications in accepted articles has achieved appropriate digital image presentation and publication as measured by a decline in corrigenda and author queries. The yearly decrease in the number of QC queries suggests that returning authors adhere to the journals' image integrity guidelines.

¹American Physiological Society, Bethesda, MD, USA, cbennett@the-aps.org

Conflict of Interest Disclosures: Rita Scheman and Christina N. Bennett are employees of the American Physiological Society, and the data presented herein are derived from American Physiological Society submissions.

Fact Checking Nucleotide Sequences in Life Science Publications: The Seek & Blastn Tool

Jennifer A. Byrne,^{1,2} Cyril Labbé³

Objective Errors within scientific publications contribute to research irreproducibility. A collection of highly similar cancer research publications (CorpusP) was recently identified, and 38 of 48 of these publications (79%) included nucleotide sequence(s) whose identities, according to blastn analyses, did not match their experimental use (either targeting an identified gene, or serving as a nontargeting control). To expand capacity to identify other studies that may incorrectly describe nucleotide sequence reagents, we aimed to design a semiautomated tool that checks the claimed use of nucleotide sequence reagents with indisputable facts from blastn homology searches; the tool was also tested with other literature claims using Google Scholar searches.

Design From a given publication, seek & blastn, a semiautomated tool, automatically extracts gene identifiers and nucleotide sequences (15 to 90 bases) using named entity recognition techniques (thesaurus and rules). The sentence containing each sequence is automatically analyzed (using finite-state machines) to assign a claimed status (targeting or nontargeting) that is compared with the most likely status according to blastn analysis. Claimed status within the literature can be further assessed by Google Scholar searches. The approach was built using the CorpusP publications and further analyzed using a set of 154 unknown studies (CorpusU) retrieved using studies from CorpusP and the "PubMed similar" functionality.

Results In CorpusP and CorpusU, 48 of 48 (100%) and 111 of 154 (73%) publications included nucleotide sequences that were extracted using seek & blastn. Application of seek & blastn identified the 38 of 48 studies (79%) in CorpusP that appear to have incorrectly employed nucleotide sequence reagent(s). More nontargeting than targeting sequences were accurately predicted to have been used incorrectly (37 of 47 [78.7%] vs 19 of 294 [6.5%]). Furthermore, the analysis of nucleotide sequences flagged by seek & blastn predicted that

30 of 154 CorpusU studies (19%) may have incorrectly employed nucleotide sequence reagent(s). However, the automated use of seek & blastn faces challenges. Overall, 10 of 341 (2.9%) and 11 of 341 (3.2%) sequences in CorpusP were either not extracted or incorrectly extracted, respectively, and claims were not (correctly) identified for 19 of 341 sequences (5.6%). Furthermore, gene identifier variations may complicate the analysis of targeting sequences. Application of seek & blastn therefore currently requires follow-up analyses by life science expert peers.

Conclusions Preliminary use of seek & blastn suggests that the incorrect use of nucleotide sequence reagents may be frequently undetected and represents an underestimated source of error in life science publications. Text mining and text analysis tools such as seek & blastn may therefore provide valuable support to allow peers to identify obvious errors in the published or forthcoming scientific literature.

¹Molecular Oncology Laboratory, Children's Cancer Research Unit, Kids Research Institute, The Children's Hospital at Westmead, Westmead, Australia; ²The University of Sydney Discipline of Child and Adolescent Health, The Children's Hospital at Westmead, Westmead, Australia; ³University of Grenoble Alpes, The National Center for Scientific Research (CNRS), Grenoble, France, cyril.labbe@imag.fr

Conflict of Interest Disclosures: Springer-Nature is funding a PhD student within the research group of Cyril Labbé. This PhD project is exploring methods to detect automatically generated scientific papers. Funding from Springer-Nature did not support the work described in this abstract.

Types of Research Integrity Issues Encountered by a Specialist Research Integrity Group

Magdalena Morawska,¹ Stephanie L. Boughton¹

Objective Data on the reasons why articles are retracted exist; however, the types and frequency of research integrity issues faced by editors day to day, particularly before publication, are unclear. Our objective was to categorize and determine the relative frequency of research integrity issues encountered by BioMed Central's Research Integrity Group, which covers approximately 300 journals spanning biological and medical disciplines.

Design We used a retrospective observational study design. We included all new inquiries regarding any aspect of research integrity sent to the Research Integrity Group between January 1, 2015, and December 31, 2016. The study period was chosen because it reflected a period when the structure and remit of the Group remained constant. The inquiries had been sent to the Research Integrity Group by editorial staff for advice and/or investigation of potential research and publication ethics issues following discussion with the journals' editors in chief. They related to submitted manuscripts or published articles and may have been detected by the editors in chief, in-house staff, peer reviewers, or whistle-blowers. Editors in chief and editorial staff are not required to escalate all issues to the Group. We assigned each inquiry to 1 of 6 categories, adapted from the Committee on

Publication Ethics (COPE) Case Taxonomy, covering different research integrity issues: authorship, competing interests, data issues, ethics/consent, peer-review process, and plagiarism/duplicate publication. Inquiries categorized as “ethics/consent” related to questions around ethics approval or consent for research involving human participants or consent for publication of potentially identifiable information (eg, case studies or images). We compared categories and their relative frequency for submitted manuscripts and published articles.

Results During the study period, the Research Integrity Group received 1040 inquiries: 690 (66%) related to submitted manuscripts and 350 (34%) to published articles. **Table 8** shows the breakdown of inquiries by category. The largest category was ethics/consent (35%), and the second largest was plagiarism/duplicate publication (23%). For inquiries relating to submitted manuscripts only, almost half (49%) related to ethics/consent. The largest category for published articles was data issues (41%). These results have been used to inform training needs for both internal staff and external editors. Editorial policies and policy wording have also been revised in line with the results of this study.

Conclusions Category frequency was different before and after publication. The high frequency of prepublication ethics/consent inquiries suggests that such issues can be detected at an early stage and that researchers need training to prevent such issues arising. Data issues were the most common for published articles, suggesting that problems with data may not always be detected by peer review and may only come to light after publication. Future studies could examine issues arising in nonbiomedical journals.

¹Springer Nature, London, UK, stephanie.boughton@biomedcentral.com

Conflict of Interest Disclosures: Both authors are employees of Springer Nature. At the time of data collection, Magdalena Morawska was an associate editor and Stephanie Boughton was a medical editor within BioMed Central’s Research Integrity Group (part of Springer Nature), the team dealing with the research and publication ethics inquiries that were the subject of this study. Both are now part of Springer Nature’s Research Integrity Group.

Acknowledgments: We thank the members of BioMed Central’s Research Integrity Group and Caroline Black for their helpful feedback on this abstract.

Table 8. Breakdown by Category of All Inquiries Received by the Research Integrity Group Between January 1, 2015, and December 31, 2016

Category	Inquiries, No (%)
Authorship	115 (11)
Competing interests	38 (4)
Data issues	182 (18)
Ethics/consent	362 (35)
Peer-review process	103 (10)
Plagiarism/duplicate publication	240 (23)

Data Sharing

Early Experiences With Journal Data Sharing Policies: A Survey of Published Clinical Trial Investigators

Sara Tannenbaum,¹ Joseph S. Ross,^{2,3,4,5} Harlan M. Krumholz,^{2,4,5,6} Nihar R. Desai,^{2,6} Jessica D. Ritchie,² Richard S. Lehman,^{2,7} Ginger M. Gamble,² Jacqueline Bachand,⁸ Sara Schroter,⁹ Trish Groves,⁹ Cary P. Gross^{3,4,10}

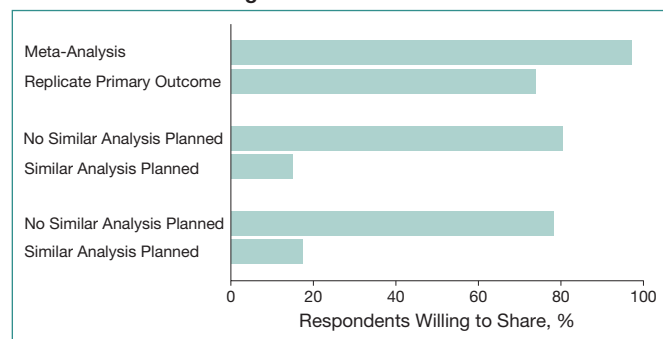
Objective Although the International Committee of Medical Journal Editors (ICMJE) recommendations for trial data sharing have been controversial, little is known about the attitudes and experiences of authors published in journals with existing data sharing policies.

Design We conducted a self-administered online survey of the authors of clinical trials published January 1, 2012, through March 1, 2016, in 3 high-impact journals with policies either requiring all clinical trial authors to share data (*PLOS Medicine*) or publish a statement specifying whether they were willing to share data (*The BMJ* and *Annals of Internal Medicine*). For the latter 2 journals, we only contacted authors who specified that they were willing to share data. We contacted the corresponding author and then an additional author if no response was received. The survey addressed sharing plans, receipt of sharing requests, and effort required to respond to sharing requests. We also asked respondents about willingness to share data in 6 hypothetical scenarios. Each hypothetical request occurred 1 year after publication of the original study but varied by type of request. Survey results are for all respondents unless otherwise indicated.

Results Among the 154 trials for which we contacted authors, 90 responses (58.4%) were received. Respondents and nonrespondents did not significantly differ by journal, year published, region of the corresponding author, or funding source. Half of the respondents had a data sharing plan (n = 49 [54.4%]), and about one-third had received at least 1 sharing request (n = 31 of 89 [34.8%]). Out of the 68 data requests that were received in aggregate, only 4 (5.9%) were denied. Most respondents indicated that they would be willing to share data for a meta-analysis (n = 87 [96.7%]) or for replication of the primary study outcome (n = 66 [73.3%]) 1 year after publication. However, in response to scenarios indicating that data were requested for a secondary outcomes analysis or predictive modeling study, willingness to share was largely influenced by author intent to conduct similar analyses (**Figure 1**). For a secondary outcomes analysis, 70 authors (77.8%) responded that they would share if they had not planned a similar analysis, but 15 authors (16.7%) responded that they would share even if they had planned a similar analysis. Among authors who had granted at least 1 request (n = 25), a median (range) of 18 (3-125) person-hours were spent to prepare data for sharing.

Conclusions Among respondents to a survey of clinical trial authors, we found that data sharing is taking place under

Figure 1. Willingness to Share by Request Type for a Trial Published 12 Months Ago



journal data sharing requirements but that willingness to share data depends on the type of request and intent to publish similar analyses.

¹Yale University School of Medicine, New Haven, CT, USA; ²Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, New Haven, CT, USA; ³Section of General Medicine, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT, USA, cary.gross@yale.edu; ⁴Robert Wood Johnson Foundation Clinical Scholars Program, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT, USA; ⁵Department of Health Policy and Management, Yale University School of Public Health, New Haven, CT, USA; ⁶Section of Cardiovascular Medicine, Yale University School of Medicine, New Haven, CT, USA; ⁷UK Cochrane Center, Oxford, UK; ⁸Yale School of Public Health, New Haven, CT, USA; ⁹The BMJ, London, UK; ¹⁰Cancer Outcomes, Public Policy, and Effectiveness Research (COPPER) Center, Yale Cancer Center and Yale University School of Medicine, New Haven, CT, USA

Conflict of Interest Disclosures: Dr Ross, Dr Krumholz, Dr Desai, Ms Ritchie, Dr Lehman, Ms Gamble, and Dr Gross receive research support from Janssen and the Pharmaceutical Companies of Johnson & Johnson, to develop methods of clinical trial data sharing. Dr Ross, Ms Ritchie, and Ms Gamble receive research support from the Blue Cross Blue Shield Association to better understand medical technology evidence generation. Drs Ross, Krumholz, and Desai receive research support from the Centers for Medicare and Medicaid Services to develop and maintain hospital performance measures that are used for public reporting. Dr Ross, Dr Krumholz, and Ms Gamble receive research support from the US Food and Drug Administration to develop methods for postmarket surveillance of medical devices. Dr Ross receives research support from the US Food and Drug Administration to establish the Yale-Mayo Center for Excellence in Regulatory Science and Innovation. Dr Gross receives research funding from 21st Century Oncology and the National Comprehensive Cancer Network-Pfizer. Dr Krumholz chairs a scientific advisory board for United Healthcare. Drs Schroter and Groves are full-time employees at *The BMJ*, and Dr Groves is a member of the Peer Review Congress Advisory Board but was not involved in the review or decision for this abstract. No other conflicts were reported.

Funding/Support: Research reported in this publication was supported by National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health (award No. 5T35DK104689-02).

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the abstract.

Disclaimer: The content is the sole responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Sharing Data Through the Yale University Open Data Access (YODA) Project: Early Experience

Joseph S. Ross,^{1,2,3,4} Jessica D. Ritchie,¹ Stephen Bamford,⁵ Jesse A. Berlin,⁵ Karla Childers,⁵ Nihar Desai,^{1,6} Ginger M. Gamble,¹ Cary P. Gross,^{2,3,7} Richard S. Lehman,^{1,8} Peter Lins,⁵ Sandra A. Morris,⁵ Joanne Waldstreicher,⁵ Harlan M. Krumholz^{1,3,4,6}

Objective To describe early experience with sharing clinical research data through the Yale University Open Data Access (YODA) Project.

Design Cross-sectional analysis of all submitted proposals by investigators to use clinical research data being made available by Johnson & Johnson (J & J) through the YODA Project, since the inception of the initiative in October 2014, including approval, data access, and publication status.

Results Of the clinical trials conducted by J & J, to date, 189 trials have been reviewed by J & J and determined to be available for sharing with external investigators, most commonly of therapies used for the treatment of bipolar disorder and schizophrenia. In addition to 1 medical device trial and 188 pharmaceutical trials, J & J continues to review trials for eligibility, and additional trials can be made available on request, including trials of consumer products. As of June 2017, the YODA Project had received 73 proposals from external investigators to use data from 159 trials; the median number of trials requested was 3 (interquartile range [IQR], 1-9; maximum, 50). Among the 73 proposals, 65 (89.0%) have been approved by an independent review panel and 2 (2.8%) are under review; 6 (8.2%) were withdrawn or closed owing to patient privacy concerns, unavailability of needed data elements, or lack of research proposal clarity. The most common study purposes proposed were to address secondary research questions (n = 39), combine data as part of larger meta-analyses (n = 35), and/or validate previously published studies (n = 17). Of the 65 approved proposals, 50 researchers have access to the data and are working on their projects (median duration of access, 43.7 weeks; IQR, 21.0-71.5), 8 are awaiting execution of their Data Use Agreement or data preparation, and 5 have completed their projects, 2 of which resulted in publications in the peer-reviewed literature, and 3 of which have submitted a manuscript for publication. In both cases, the final publication represented the originally proposed research. The authors of the remaining 2 proposals did not pursue their projects.

Conclusions Early experience sharing data through the YODA Project has demonstrated a demand for shared clinical research data as a resource for investigators. As trial funders and investigators increasingly share data, and make use of shared data, it is essential to understand best practices and incentives to ensure success.

¹Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, New Haven, CT, USA, joseph.ross@yale.edu; ²Section of General Internal Medicine, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT, USA; ³Robert

Wood Johnson Foundation Clinical Scholars Program, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT, USA; ⁴Department of Health Policy and Management, Yale School of Public Health, New Haven, CT, USA; ⁵Johnson & Johnson, New Brunswick, NJ, USA; ⁶Section of Cardiovascular Medicine, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT, USA; ⁷Cancer Outcomes, Public Policy, and Effectiveness Research (COPPER) Center, Yale Cancer Center and Yale University School of Medicine, New Haven, CT; ⁸UK Cochrane Center, Oxford, UK

Conflict of Interest Disclosures: Drs Ross, Desai, Gross, Lehman, and Krumholz and Mss Ritchie and Gamble receive research support from Janssen, the Pharmaceutical Companies of Johnson & Johnson, to develop methods of clinical trial data sharing. Dr Ross and Mss Ritchie and Gamble receive research support from the Blue Cross Blue Shield Association (BCBSA) to better understand medical technology evidence generation. Drs Desai, Ross, and Krumholz receive research support from the Centers for Medicare and Medicaid Services to develop and maintain hospital performance measures that are used for public reporting. Drs Ross and Krumholz and Ms Gamble receive research support from the US Food and Drug Administration (FDA) to develop methods for post-market surveillance of medical devices. Dr Ross and Ms Ritchie receive research support from the FDA to establish the Yale-Mayo Center for Excellence in Regulatory Science and Innovation. Dr Gross receives research funding from 21st Century Oncology and the National Comprehensive Cancer Network–Pfizer. Dr Krumholz chairs a scientific advisory board for United Healthcare. Drs Berlin, Morris, and Waldstreicher, Ms Childers, and Messrs Bamford and Lins are employees of Johnson & Johnson.

Funding/Support: This research was the direct result of experience gained from a research agreement through Yale University from Johnson & Johnson (Janssen) to develop methods of clinical trial data sharing.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the abstract.

Disclaimer: The authors assume full responsibility for the accuracy and completeness of the ideas presented.

Statements About Intent to Share Individual Participant Data at ClinicalTrials.gov

Annice Bergeris,¹ Tony Tse,¹ Deborah A. Zarin¹

Objective Following recent calls for clinical trialists to declare their plans to share individual participant data (IPD)

prior to study initiation, ClinicalTrials.gov added 2 optional data elements in December 2015: (1) Plan to Share IPD (submitted at study initiation) and (2) Available Study Data/Documents (submitted after study completion). We sought to characterize responses to ClinicalTrials.gov’s IPD sharing–related questions.

Design We summarized responses to IPD sharing–related questions for all interventional studies initially registered in 2016 and for the subset of trials registered by “high-volume data providers,” defined as organizations that registered 10 or more records. Organizations were categorized by key funder type, ie, were categorized as “NIH” if at least 1 National Institutes of Health institute was listed as a sponsor or collaborator, as “industry” if not classified as NIH and at least 1 company was listed as a sponsor or collaborator, and as “other” for all remaining records. Because of considerable heterogeneity in responses among the high-volume data provider subsample, we further characterized responses for trials registered by the top 10 high-volume data provider organizations within each key funder type.

Results Of 21,310 trial records analyzed by May 10, 2017, 14,523 (68.2%) included a response to the question about plans to share IPD; 1930 records (13.3%) indicated yes, 3821 (26.3%) indicated undecided, and 8772 (60.4%) indicated no. Proportions within each key funder type varied among the 10,894 records from high-volume data providers (**Table 9**). Among the top 10 organizations within each key funder type, the percentage of records indicating that plans exist for sharing IPD ranged from 0% to 24% for NIH-funded studies, 0% to 63% for industry-funded studies, and 0% to 28% of studies with other funding. Among 131 records indicating that documents were available for sharing, 76 specified the study protocol would be shared, 52 specified other documents (eg, information leaflets), 32 specified informed consent forms, 16 specified the clinical study report, and 14 specified the individual participant data set. Five of 14 records (36%) specifying availability of IPD listed no plans for sharing IPD.

Conclusions Sixty-eight percent of trial registrants responded to an optional question about plans to share IPD. Among those respondents, 13% said they would share data and another 26% were undecided. Of the 131 records indicating availability of documents for sharing, only 14

Table 9. Responses to Trial Registration Question About Plan to Share Individual Participant Data Among High-Volume Data Providers^a

Summary Statistics	Total	Key Funder Type		
		NIH	Industry	Other
Organizations (trial records), No.	429 (10,894)	23 (484)	69 (2031)	337 (8379)
≥1 Response (records, No. [%])	404 (6802 [62.4])	18 (178 [36.8])	52 (566 [27.9])	335 (6058 [72.3])
% Yes, median (range)	6 (0-92)	8 (0-36)	0 (0-88)	7 (0-92)
% Undecided, median (range)	14 (0-100)	6 (0-64)	0 (0-100)	17 (0-87)
% No, median (range)	43 (0-100)	18 (0-69)	6 (0-100)	45 (0-100)
≥1 Missing response (records, No. [%])	401 (4092 [37.6])	22 (306 [63.2])	63 (1465 [72.1])	316 (2321 [27.7])
% No response, median (range)	25 (0-100)	57 (0-100)	73 (0-100)	23 (0-100)

Abbreviation: NIH, National Institutes of Health.
^aHigh-volume data providers defined as organizations that registered 10 or more trial records.

indicated that IPD were available. Considerable cultural and scientific changes will be necessary before the sharing of IPD and associated documents becomes part of routine practice by clinical researchers.

¹ClinicalTrials.gov, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA, dzarin@mail.nih.gov

Conflict of Interest Disclosures: All authors work for ClinicalTrials.gov. Dr Zarin is a member of the Peer Review Congress Advisory Board but was not involved in the review or decision for this abstract.

Funding/Support: Supported by the Intramural Research Program of the National Library of Medicine, National Institutes of Health.

Monday, September 11, 2017

Improving Peer Review and Scientific Publication

Quality of Reporting

Association of Journal-Level and Study-Level Variables With Proximity of Primary Study Results to Summary Estimates From Meta-analyses in Imaging Journals

Robert Frank,¹ Matthew McInnes,^{2,3} Deborah Levine,^{4,5} Herbert Kressel,^{4,5} Julie Jesurum,⁶ William Petreich,³ Trevor McGrath,¹ Patrick M. Bossuyt⁷

Objective Are most research findings false but less so in high-impact journals? Variables associated with false results in imaging research are poorly understood. While absolute truth is elusive, high-quality meta-analyses using hierarchical methods produce high-level evidence with minimal random variability in their results, thereby representing a convenient proxy for truth. We evaluated whether journal-level and study-level variables are associated with the distance between primary study results and summary estimates from meta-analyses.

Design In this meta-research study, PubMed was searched for diagnostic accuracy meta-analyses using hierarchical methods, published in imaging journals between January 2005 and April 2016. Data were extracted for each meta-analysis and its included primary studies, including study demographic information, journal Impact Factor, journal cited half-life, Standards for Reporting Diagnostic accuracy studies (STARD) endorsement, citation rate, publication date, sample size, sensitivity, and specificity. Meta-analyses were excluded for failing to report both primary and summary accuracy estimates. Primary studies were divided into 2 groups for each variable assessed; groups were defined based on first publication vs subsequent publications on a topic, publication before vs after STARD introduction, presence vs absence of STARD endorsement, or by median split. The mean absolute deviation of primary study estimates from the corresponding summary estimates for sensitivity and specificity was compared between groups for each variable.

Analyses were performed using a model combining a γ distribution for absolute deviations greater than 0 with an estimated probability that the absolute deviation is 0. Means and 95% CIs were obtained using bootstrap resampling. *P* values were calculated using a *t* test. The threshold for significance was defined as $P < .004$ after Bonferroni correction ($.05/12$) to mitigate bias owing to multiple comparisons.

Results Ninety-eight meta-analyses containing 1458 primary studies met inclusion criteria. There was substantial variability in deviations from the summary estimate between paired groups, but no variable demonstrated a significant association with proximity of primary study diagnostic accuracy estimates to the pooled estimates from their corresponding meta-analyses ($P > .004$ in all comparisons) (Table 10).

Conclusions Many variables considered important when selecting imaging diagnostic accuracy literature to guide clinical decisions are not associated with results that are more reflective of the truth as established by meta-analyses. The distance between primary study results and summary estimates of diagnostic accuracy is probably not smaller for studies published in higher versus lower Impact Factor journals.

¹Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada, robert.frank@uottawa.ca; ²Department of Radiology, The Ottawa Hospital, University of Ottawa, Ottawa, ON, Canada; ³Ottawa Hospital Research Institute, Ottawa, ON, Canada; ⁴Department of Radiology, Beth Israel Deaconess Medical Center, Boston, MA, USA; ⁵Radiology Editorial Office, Boston, MA, USA; ⁶Harvard University, Boston, MA, USA; ⁷Clinical Epidemiology and Biostatistics,

Table 10. Association of Journal-Level and Study-Level Variables With Proximity of Primary Study Results to Summary Estimates From Meta-analyses^a

Variable	Group Analysis	Difference in Mean Deviation of Primary Estimates From Summary Estimates Between Dichotomized Groups			
		Sensitivity	<i>P</i> Value	Specificity	<i>P</i> Value
Impact Factor	Above median vs below median	−0.018	.09	−0.013	.11
STARD endorsement	Endorsement vs no endorsement	−0.0057	.60	0.0019	.83
Cited half-life	Above median vs below median	−0.0063	.55	0.0097	.24
Citation rate	Above median vs below median	−0.018	.08	−0.0045	.58
Publication timing (relative to STARD 2003)	Post-STARD vs pre-STARD	0.0059	.55	−0.0082	.38
Publication timing (first published)	First published vs later published	−0.025	.005	0.0077	.48

Abbreviation: STARD, Standards for Reporting Diagnostic accuracy studies.

^aStatistical significance defined as $P < .004$ (after Bonferroni correction).

Conflict of Interest Disclosures: None reported.

Funding/Support: This work was supported in part by stipends from the Ottawa Hospital Department of Radiology and the Undergraduate Research Opportunity Program.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the abstract.

Discrepancies in Reporting Between Trial Publications and Clinical Trial Registries in High-Impact Journals

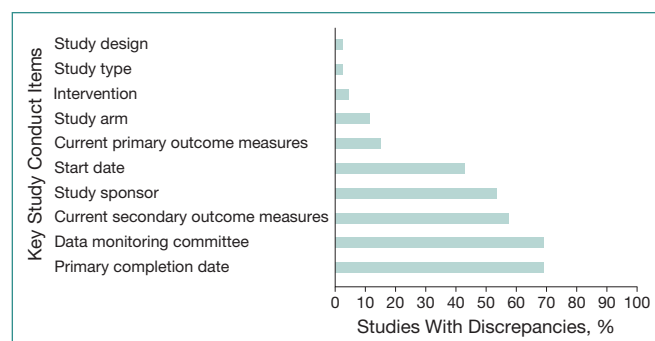
Sarah Daisy Kosa,^{1,2} Lawrence Mbuagbaw,¹ Victoria Borg Debono,¹ Mohit Bhandari,¹ Brittany B. Dennis,¹ Gabrielle Ene,⁴ Alvin Leenus,³ Daniel Shi,³ Michael Thabane,¹ Thuva Vanniyasingam,¹ Chenglin Ye,⁵ Elgene Yranon,¹ Shiyuan Zhang,¹ Lehana Thabane¹

Objective It is currently unclear the extent to which key information mandatory for clinical trial registries is reported in published manuscripts. To address this gap in the literature, the primary objective of this study was to examine the percentage of studies where there are discrepancies in reporting of key study conduct items between the clinical trial registry and the manuscript.

Design We searched PubMed for all randomized clinical trials (RCTs) published between 2012 and 2015 in the top 5 general medicine journals (based on the 2014 impact factor as published by Thomson Reuters), which all required registration of the RCT for publication; 200 full-text publications (50 from each year) were randomly selected for data extraction. Key study conduct items were extracted by 2 independent reviewers for each year. When an item was reported differently or not reported at all in either source, this was considered a discrepancy in reporting between the registry and the full-text publication. Descriptive statistics were calculated to summarize the percentage of studies with discrepancies between the registry and the published manuscript in reporting of key study conduct items. The items of interest were design (ie, randomized control trial, cohort study, case control study, case series), type (ie, retrospective, prospective), intervention, arms, start and end dates (based on month and year where available), use of data monitoring committee, and sponsor, as well as primary and secondary outcome measures.

Results In the sample of 200 RCTs, there were relatively few studies with discrepancies in study design (n=6 [3%]), study type (n=6 [3%]), intervention (n=10 [5%]), and study arm (n=24 [12%]) (Figure 2). Only 30 studies (15%) had discrepancies in their primary outcomes. However, there were often discrepancies in study start date (n=86 [43%]), study sponsor (n=108 [54%]), and secondary outcome measures (n=116 [58%]). Almost 70% of studies had

Figure 2. Percentage of Studies With Discrepancies in Reporting Between Trial Publications and Clinical Trial Registries in 200 High-Impact Journals



discrepancies regarding the use of a data monitoring committee and primary completion date reporting.

Conclusions We identified discrepancies in reporting between publications and clinical trial registries. These findings are limited by only being based on a subset of RCTs in the included journals and may not be generalizable to all RCTs within that journal, other disciplines, journals in other languages, or lower-impact journals. Further measures are needed to improve reporting given the potential threats to the quality and integrity of scientific research.

¹Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada, mbuagblc@mcmaster.ca; ²Toronto General Hospital-University Health Network, Toronto, Ontario, Canada; ³Princess Margaret Hospital-University Health Network, Toronto, Ontario, Canada; ⁴Bachelor of Health Sciences Program, McMaster University, Hamilton, Ontario, Canada; ⁵Oncology Biostatistics, Genentech, South San Francisco, CA, USA

Conflict of Interest Disclosures: None reported.

Methodological and Reporting Quality of Systematic Reviews Underpinning Clinical Practice Guidelines

Cole Wayant,¹ Matt Vassar¹

Objective This study summarizes the findings of 3 separate studies conducted simultaneously to determine the methodological and reporting quality of systematic reviews (SRs) underpinning clinical practice guidelines (CPGs) in pediatric obesity, opioid use disorder, and ST-elevated myocardial infarction.

Design A search of guideline clearinghouse and professional organization websites was conducted for guidelines published by national or professional organizations. We included all reviews cited by authors of CPG, including Cochrane reviews, and removed duplicates prior to data extraction. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) and AMSTAR (A Measurement Tool to Assess Systematic Reviews) instruments were used to score SRs and meta-analyses cited in CPGs. PRISMA and AMSTAR are validated tools for measuring reporting quality and methodological quality, respectively.

Results The mean PRISMA total scores for the pediatric obesity, opioid use disorder, and ST-elevated myocardial infarction SRs across all CPGs were 16.9, 20.8, and 20.8, respectively. The mean AMSTAR total scores were 4.4, 8.8, and 6.1, respectively. Consistently underreported items on the PRISMA checklist were items 5 (protocol registration), 8 (search strategy), 15 (risk of bias for cumulative evidence), and 22 (risk of bias across studies). Consistently underreported items on the AMSTAR checklist were items 4 (duplicate extraction/validation), 5 (list of included/excluded studies), 8 (quality of evidence assessments), 10 (publication bias assessments), and 11 (conflict of interest disclosure). Altogether, our study included 150 SRs and 29 CPGs, with only 9 CPGs assigning grades to their recommendations. The 150 SRs were cited a total of 308 times: 95 times as a direct evidence for graded recommendations, 21 times as direct evidence for nongraded recommendations, 189 times as supporting evidence, and 3 times for unclear reasons.

Conclusions These investigations into CPGs in pediatric obesity, opioid use disorder, and ST-elevated myocardial infarction revealed a consistent lack of overall methodological and reporting quality in the included SRs as well as heterogeneity in the use of grading scales, or lack thereof. Because SRs are considered by most to be level 1A evidence, an apparent lack of quality may impair clinical decision making and hinder the practice of evidence-based medicine. Items such as PRISMA items 15 and 22 and AMSTAR items 10 and 11 are of particular concern because these items ensure that bias assessments are performed and conflicts of interests are disclosed.

¹Oklahoma State University Center for Health Sciences, Department of Analytical and Institutional Research, Tulsa, OK, USA, cole.wayant@okstate.edu

Conflict of Interest Disclosures: None reported.

Additional Contributors: We acknowledge the work done by the members of the team that conducted the investigations into opioid use disorder (Andrew Ross and Justin Rankin) and ST-elevated myocardial infarction (Jared Scott and Ben Howard) as well as all of the faculty and staff at the Oklahoma State University Center for Health Sciences and Oklahoma State University Medical Center that assisted in the process of completing the investigations.

Optimism Bias in Contemporary National Clinical Trial Network Phase 3 Trials

Kaveh Zakeri,¹ Sonal S. Noticewala,¹ Lucas K. Vitzthum,¹ Elena Sojourner,¹ Loren K. Mell¹

Objective Overestimation of treatment effect sizes—termed *optimism bias*—in research protocols can lead to underpowered clinical trials that fail to demonstrate clinical benefits. We compared hypothesized vs. observed treatment effects to determine if there is evidence of optimism bias in contemporary NCTN phase III trials.

Design We queried PubMed for National Cancer Institute (NCI)—sponsored phase III randomized cooperative group clinical trials from January 2007 to January 2017. We

identified 185 published trials. Trials with missing protocols ($n = 56$), equivalence or noninferiority trials ($n = 5$), trials that accrued less than 40% of their intended sample size ($n = 14$), and trials that pooled their data with other studies ($n = 2$) were excluded. For trials reporting time-to-event outcomes with hazard ratios (HRs) ($n = 81$), we compared the proposed effect size from the sample size calculation in the research protocol with the observed effect size in the published article to calculate the ratio of observed-to-proposed HRs overall and for trials that did or did not report statistically significant effect on primary end points. All HRs were standardized for a reduction in adverse events such that HRs less than 1 indicated a benefit to therapy. We also compared findings with those previously reported for NCI trials conducted from 1955 to 2006 and tabulated studies that provided a reference, evidence, or other specific rationale for their proposed effect size in the research protocol.

Results Data on 98,200 patients from 108 clinical trials were evaluated. The most common cancers were breast, gynecologic, gastrointestinal, brain, and genitourinary malignant neoplasms. The most common primary end point was overall survival (40.7%). The median ratio of observed-to-proposed HRs was 1.26 (range: 0.33–2.34). The median ratio of observed-to-proposed HRs among trials that observed a statistically significant effect on the primary end point was 1.09 (range: 0.33–1.29) vs 1.30 (range: 0.86–2.34) for trials that did not, compared with 1.34 and 1.86, respectively, for NCI trials conducted from 1955 to 2006. Twenty-four trials (22.2%) observed a statistically significant effect on the primary end point favoring the experimental treatment, compared with 24.6% previously reported. The majority of trials (76.9%) provided no rationale for the magnitude of the proposed treatment effect.

Conclusions Although most NCI-sponsored clinical trials conducted between 2007 and 2017 failed to establish statistically significant benefits of new therapies, the magnitude of optimism bias appears to have decreased compared with that in trials conducted between 1955 and 2006. Better rationalization of proposed effect sizes is needed in clinical trial protocols.

¹Department of Radiation Medicine and Applied Sciences, University of California, San Diego, La Jolla, CA, USA, kzakeri@ucsd.edu

Conflict of Interest Disclosures: None reported.

Quality of the Scientific Literature

Scientific Quality in a Series of Comparative Effectiveness Research Studies

Harold Sox,^{1,2} Evan Mayo-Wilson,^{1,2} Kelly Vander Ley,^{1,2} Marina Broitman,^{1,2} David Hickam,^{1,2} Steven Clauser,^{1,2} Yen-Pin Chiang,^{1,2} Evelyn Whitlock^{1,2}

Objective Markers of high-quality comparative effectiveness research (CER) studies are largely unknown but could be valuable to funders and future applicants for CER funding.

Our long-term objective is to identify variables associated with CER scientific quality and impact. The objective of this preliminary report is to describe the frequency of measures of CER study quality.

Design This is a case series of CER studies funded during the first funding cycle (2013) of the Patient-Centered Outcomes Research Institute (PCORI). Awardees are required to submit a final research report (FRR), which undergoes external peer review and is published on the PCORI website when the principal investigator (PI) meets revision requirements. We are using the original application to investigate study and PI-related variables potentially associated with study quality, and are assessing study quality of the peer-reviewed report using US Preventive Services Task Force (USPSTF) criteria (good, fair, poor) and adherence to PCORI methodology standards. When the case series is complete we will study associations between markers of study quality and publication outcomes (citations in published articles, systematic reviews, and practice guidelines; Altmetric scores) and between direct study quality measures and those outcomes.

Results Among 98 FRRs received by early June 2017, 5 have completed peer review. Candidate PI-based variables potentially associated with study quality include number of research awards from the National Institutes of Health, Agency for Healthcare Research and Quality, Centers for Disease Control and Prevention, or the Department of Veterans Affairs (mean, 1 [range, 0-5]; median, 0); number of studies in major journals as first or last author (mean, 2.8 [0-8]; median, 0), the PI's H factor (mean, 23 [range, 16-41]; median, 22), and years since he or she was granted the highest academic degree (19.2 [range, 16-21]; median, 20). All 5 studies were of "fair" quality according to USPSTF grading criteria. Each of PCORI's 5 cross-cutting Methodology Standards (which had not been published when the studies in this report were funded) comprise several component standards (range, 7-17), and rates of meeting the standards varied from 15% (standard for managing missing data) to 34% (standard for formulating research questions). We expect to complete peer review and to report on 20 more research reports.

Conclusions With short-term follow-up on a series of approximately 300 CER studies funded by PCORI through July 2019, this study may eventually provide measures of specific methodological shortcomings and variables associated with CER quality and impact.

¹Patient-Centered Outcomes Research Institute, hsox@pcori.org, Johns Hopkins University, Baltimore, MD, USA; ²Oregon Health Sciences University, Portland, OR, USA

Conflict of Interest Disclosures: None reported.

Pitfalls in the Use of Statistical Methods in Systematic Reviews of Therapeutic Interventions: A Cross-sectional Study

Matthew J. Page,¹ Douglas G. Altman,² Larissa Shamseer,^{3,4} Joanne E. McKenzie,¹ Nadera Ahmadzai,⁵ Dianna Wolfe,⁵ Fatemeh Yazdi,⁵ Ferrán Catalá-López,^{5,6} Andrea C. Tricco,^{7,8} David Moher^{3,4}

Objective Researchers have identified several problems in the application of statistical methods in published systematic reviews (SRs). However, these evaluations have been narrow in scope, focusing only on one particular method (such as sensitivity analyses) or restricting inclusion to Cochrane SRs, which make up only 15% of all SRs of biomedical research. We aimed to investigate the application and interpretation of various statistical methods in a cross-section of SRs of therapeutic interventions, without restriction by journal, clinical condition, or specialty.

Design We selected articles from a database of SRs we assembled previously. These articles consisted of a random sample of 300 SRs addressing various questions (therapeutic, diagnostic, or etiologic) that were indexed in MEDLINE in February 2014. In the current study, we included only those SRs that focused on a therapeutic question, reported at least 1 meta-analysis, and were written in English. We collected data on 61 prespecified items that characterized how well random-effects meta-analysis models, subgroup analyses, sensitivity analyses, and funnel plots were applied and interpreted. Data were extracted from articles and online appendices by a single reviewer, with a 20% random sample extracted in duplicate.

Results Among 110 SRs, 78 (71%) were non-Cochrane SRs and 55 (50%) investigated a pharmacological intervention. The SRs presented a median of 13 (interquartile range, 5-27) meta-analyses. Among the 110 primary meta-analyses in each SR, 62 (56%) used the random-effects model but only 5 of 62 (8%) interpreted the pooled result correctly (that is, as the average of the intervention effects across all studies). Subgroup analyses were reported in 42 of 110 SRs (38%), but findings were not interpreted with respect to a test for interaction in 29 of 42 cases (69%), and the issue of potential confounding in the subgroup analyses was not raised in any SR. Sensitivity analyses were reported in 51 of 110 SRs (46%), without any rationale in 37 of 51 cases (73%). Authors of 37 of 110 SRs (34%) reported that visual inspection of a funnel plot led them to not suspect publication bias. However, in 28 of 37 cases (76%), fewer than 10 studies of varying size were included in the plot.

Conclusions There is scope for improvement in the application and interpretation of statistical analyses in SRs of therapeutic interventions. Guidelines such as PRISMA may need to be extended to provide more specific statistical guidance.

¹School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia, matthew.page@monash.edu; ²UK EQUATOR Centre, Centre for Statistics in Medicine, NDORMS, University of Oxford, Oxford, UK; ³Centre

for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada; ⁴School of Epidemiology, Public Health, and Preventive Medicine, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada; ⁵Knowledge Synthesis Group, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada; ⁶Department of Medicine, University of Valencia/INCLIVA Health Research Institute and CIBERSAM, Valencia, Spain; ⁷Knowledge Translation Program, Li Ka Shing Knowledge Institute, St Michael's Hospital, Toronto, Ontario, Canada; ⁸Epidemiology Division, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada

Conflict of Interest Disclosures: Douglas G. Altman and David Moher are Peer Review Congress Advisory Board Members but were not involved in the review or decision for this abstract.

Introducing Reporting Guidelines and Checklists for Contributors to *Radiology*: Results of an Author and Reviewer Survey

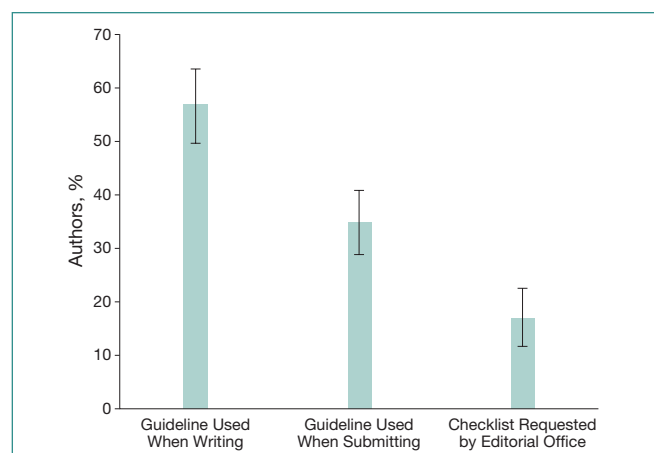
Marc Dewey,¹ Deborah Levine,^{2,3} Patrick M. Bossuyt,⁴ Herbert Y. Kressel^{5,6}

Objective Numerous reporting guidelines have been developed to make study reports more informative, but it is uncertain whether they are perceived as useful by authors and reviewers. We surveyed the use and perceived value of reporting guidelines after an initiative begun in January 2016 that required authors to submit appropriate guideline checklists along with their manuscripts prior to peer review.

Design Cohort study of authors of original research submissions to *Radiology* between July 5, 2016, and June 1, 2017, and of reviewers who had performed reviews since January 2016. Authors were asked to complete an anonymized online survey within 2 weeks of manuscript submission but before the editorial decision was made. Reviewers were surveyed with similar questions from May 17, 2017, until June 1, 2017.

Results A total of 831 of 1391 authors (59.7%) completed the survey within a mean (SD) of 1.5 (2.7) days (range, 0–17 days) of the request. Consistent with the types of studies submitted to *Radiology*, most authors used STROBE (447 of 829 authors [53.9%]) or STARD (313 authors [37.8%]) and only a small minority used CONSORT (40 authors [4.8%]) or PRISMA (29 authors [3.5%]). Only 120 of 821 authors (14.6%) used the guideline and checklist when designing the study, more so for PRISMA users (16 of 29 [55%]), less so for STARD users (52 of 310 [16.8%]; $P < .001$) and STROBE users (46 of 443 [10.4%]; $P < .001$). The guidelines were used by 189 of 821 authors (23.0%) when writing the manuscript; these authors more often reported an impact on the final manuscript (107 of 189 [56.6%]) compared with those who used the guideline when submitting the manuscript (95 of 272 [34.9%]; $P < .001$) or when the checklist was requested by the editorial office (41 of 240 [17.1%]; $P < .001$) (Figure 3). Filling out the checklist was considered very useful by 256 of 819 authors (31.3%), somewhat useful by 390 (47.6%), not very useful by 122 (14.9%), and not at all useful by 51 (6.2%). The response rate of reviewers was 32.1% (259 of 808 reviewers). The checklist was used by 200 of 259

Figure 3. Impact on Manuscripts Depending on When Guideline and Checklist Were Used



Significantly more authors who used the guidelines and checklist when writing the manuscript reported an impact on the final manuscript (107 of 189 authors [56.6%]) compared with those who used the guideline when submitting the manuscript (95 of 272 authors [34.9%]; $P < .001$) or when the checklist was requested by the editorial office (41 of 240 authors [17.1%]; $P < .001$). Error bars show 95% CIs.

reviewers (77.2%) some or all of the time, and 119 of 199 (59.8%) said it affected their reviews. Having the checklist for review was considered very useful by 28 of 256 reviewers (10.9%), somewhat useful by 106 (41.4%), not very useful by 82 (32.0%), and not at all useful by 40 (15.6%).

Conclusions Almost 4 of 5 authors and half of the reviewers judged the guideline checklists to be useful or very useful. Using the guidelines while writing the manuscript was associated with greater impact on the final manuscript.

¹Department of Radiology, Charité—Universitätsmedizin Berlin, Berlin, Germany, marc.dewey@charite.de; ²Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA; ³Senior Deputy Editor, *Radiology*, Boston, MA, USA; ⁴Amsterdam Public Health Research Institute, Department of Clinical Epidemiology, Biostatistics, and Bioinformatics, University of Amsterdam, Amsterdam, the Netherlands; ⁵Department of Radiology, Harvard Medical School, Boston, MA, USA; ⁶Editor, *Radiology*, Boston, MA, USA

Conflict of Interest Disclosures: Marc Dewey was an Associate Editor of *Radiology* at the time of initiation of this study and is a consultant to the editor now. Debbie Levine is Senior Deputy Editor of *Radiology*. Patrick Bossuyt is the lead senior author of the STARD guidelines and checklist. Herbert Y. Kressel is the Editor in Chief of *Radiology* and an author of the STARD 2015 guidelines.

Funding/Support: This study was supported by the Young Leaders Club program of the International Society of Strategic Studies in Radiology and the Heisenberg Program of the German Research Foundation.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the abstract.

Reported Use of Standard Reporting Guidelines Among JNCI Authors, Editorial Outcome, and Reviewer Ratings Related to Adherence to Guidelines and Clarity of Presentation

Jeannine Botos¹

Objective A study was conducted to examine associations between author-reported use of standard reporting guidelines (SRGs) to prepare JNCI submissions with editorial decisions and reviewer ratings for adherence to reporting guidelines and clarity of presentation.

Design At submission authors were asked if they used SRGs to prepare their manuscript and, if so, which one(s). Reviewers rated (poor, fair, good, very good, outstanding, not applicable) adherence to reporting guidelines and clarity of presentation. This information was collected using a customized Editorial Manager Enterprise Analytics Report for submissions with first or final decisions that were submitted between November 1, 2015 and April 30, 2017. All manuscript types that would benefit from the use of SRGs were included (ie, Articles, Brief Communications, Reviews, MiniReviews, Systematic Reviews, and Meta-analyses). Each peer-reviewed submission received 1 to 3 ratings per question and all ratings were included in the analyses. Numerical values were given to each answer (SRG use, 1; no SRG use, 0) or reviewer rating (not applicable, 0; fair, 1; poor, 2; good, 3;

very good, 4; and outstanding, 5), and scores were compared using 2-sided *t* tests.

Results Of 2209 submissions included in the analysis, 1144 (51.8%) indicated that at least 1 SRG was used (Table 11). The STROBE guidelines were the most common (*n* = 531, 24.0%). Of the 2068 (93.6%) submissions that were rejected, 1105 (50.1%) indicated using SRGs and 963 (43.6%) did not (mean [SD] scores of rejected vs not rejected, 0.53 [0.50] vs 0.49 [0.50], *P* = .47). Of the 1033 ratings for adherence to reporting guidelines, mean (SD) scores for not rejected vs rejected submissions were 3.2 (1.61) vs 2.9 (1.57) (*P* = .005), and mean (SD) scores for SRG use vs no use were 3.1 (1.48) vs 2.9 (1.70) (*P* = .01). Of the 1036 ratings for clarity of presentation, mean (SD) scores for not rejected vs rejected submissions were 3.6 (1.00) vs 3.1 (1.08) (*P* < .001), whereas mean (SD) scores for SRG use vs no use were 3.3 (1.04) vs 3.3 (1.10) (*P* = .64).

Conclusions Among these JNCI submissions, reporting the use of SRGs was not associated with editorial decisions or with reviewer ratings for clarity of presentation. Reviewer ratings for adherence to guidelines and clarity of presentation were associated with editorial decisions after peer review, and ratings for adherence to guidelines were associated with reported use of SRGs.

¹JNCI, Oxford University Press, New York, NY, USA, jeannine.botos@oup.com

Conflict of Interest Disclosures: None reported.

Table 11. Reported Use of Standard Reporting Guidelines Among JNCI Authors, Editorial Outcomes, and Reviewer Ratings for Adherence to Guidelines and Clarity of Presentation for Articles, Reviews, Mini-Reviews, Systematic Reviews, Meta-analysis, and Brief Communications^a

Editorial Decision or Reviewer Question	All, no (%)	Reported Using a SRG				Adherence to Reporting Guidelines					Clarity of Presentation				
		No, no. (%)	Any, no. (%)	Mean Score (SD)	<i>P</i> ^b	All, no. (%)	No SRG, no. (%)	Any SRG, no. (%)	Mean Score (SD)	<i>P</i> ^c	All, no. (%)	No SRG, no. (%)	Any SRG, no. (%)	Mean Score (SD)	<i>P</i> ^c
All submissions	2209 (100)	1065 (48.2)	1144 (51.8)	0.52 (0.5)		1033 (100)	552 (53.4)	481 (46.6)	3.0 (1.6)		1036 (100)	487 (47.0)	549 (53.0)	3.3 (1.1)	
Rejected without peer review	1813 (82.1)	875 (39.6)	938 (42.5)	0.53 (0.5)		NA	NA	NA	NA		NA	NA	NA	NA	
Rejected after peer review	255 (11.5)	88 (4.0)	167 (7.6)	0.53 (0.5)	.68	609 (58.9)	343 (33.2)	266 (25.7)	2.9 (1.6)		608 (58.7)	340 (32.8)	268 (27.6)	3.1 (1.1)	
Not rejected after peer review	141 (6.4)	102 (4.6)	39 (1.8)	0.49 (0.5)	.47	424 (41.0)	209 (20.2)	215 (20.8)	3.2 (1.6)	.004	428 (41.3)	219 (21.1)	209 (20.2)	3.6 (1.0)	<i>P</i> < .001
Adherence to reporting guidelines, mean score (SD)	3.0 (1.6)	2.9 (1.7)	3.1 (1.5)		.01										
Clarity of presentation, mean score (SD)	3.3 (1.1)	3.3 (1.1)	3.3 (1.0)		.64										

Abbreviations: SRG, standard reporting guideline; NA, not applicable.

^aAuthors reported using the following SRGs: Strengthening-Reporting of Observational-Studies in Epidemiology (STROBE); Animal Research: Reporting In Vivo Experiments (ARRIVE); Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE); Consolidated Standards of Reporting Trials (CONSORT); REporting recommendations for tumour MARKer prognostic studies (REMARK); Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA); studies of diagnostic accuracy (STARD); Meta-analyses of Observational Studies (MOOSE); Biospecimen reporting for improved study quality (BRISQ); STrengthening the REporting of Genetic Association Studies (STREGA), an extension to STROBE; and Consolidated Health Economic Evaluation Reporting Standards (CHEERS). Some percentages do not add to 100 owing to rounding. Numerical values were given to each answer (SRG use, 1; no SRG use, 0) or reviewer rating (not applicable, 0; fair, 1; poor, 2; good, 3; very good, 4; and outstanding, 5), and mean scores are presented. *P* values were calculated using a 2-sided paired *t* test.

^b*P* comparing scores for SRG use vs no SRG use.

^c*P* comparing scores for rejected vs not rejected editorial decisions.

Impact of an Intervention to Improve Compliance With the ARRIVE Guidelines for the Reporting of In Vivo Animal Research

Emily Sena,¹ for the Intervention to Improve Compliance With the ARRIVE Guidelines (IICARus) Collaborative Group

Objective To conduct a randomized controlled trial to determine whether journal-mandated completion of an ARRIVE checklist (requiring authors to state on which page of their manuscript each checklist item is met) improves full compliance with the ARRIVE guidelines.

Design Manuscripts submitted to *PLOS One* between March 2015 and June 2015 determined in the initial screening process to describe in vivo animal research were randomized to either mandatory completion and submission of an ARRIVE checklist or the normal editorial processes, which do not require any checklist submission. The primary outcome was between-group differences in the proportion of studies that comply with the ARRIVE guidelines. We used online randomization with minimization (weighted at 0.75) according to country of origin; this was performed by the journal during technical checks after submission. Authors, academic editors, and peer reviewers were blinded to the study and the allocation. Accepted manuscripts were redacted for information relating to the ARRIVE checklist by an investigator who played no further role in the study to ensure outcome adjudicators were blinded to group allocation. We performed outcome adjudication in duplicate by assessing manuscripts against an operationalized version of the ARRIVE guidelines that consists of 108 items. Discrepancies are being resolved by a third independent reviewer.

Results We randomly assigned 1689 manuscripts, with 844 manuscripts assigned to the control arm and 845 assigned to the intervention arm. Of these, 1299 (76.9%) were sent for review, and of these, 688 (53.0%) were accepted for publication. All 688 manuscripts were dual assessed, and reconciliation of discrepancies is ongoing. Agreement between reviewers was high in relation to questions of the species reported (93%) and measures to reduce the risk of bias (73%–91% for 6 questions) and lowest for reporting the unit of analysis (50%). Data analysis is ongoing. We will present data for between-group differences in the proportion of studies that comply with the ARRIVE guidelines, each of the 38 subcomponents of the ARRIVE checklist, each of the 108 items, and the proportion of submitted manuscripts accepted for publication.

Conclusions Our study will determine the effect of an alteration of editorial policy to include a completed ARRIVE checklist with submissions on compliance with the ARRIVE guidelines in the work when published. These results will inform the future development and further implementation of the ARRIVE guidelines.

¹Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK, emily.sena@ed.ac.uk

Conflict of Interest Disclosures: The study management committee included a representative from the Public Library of Science (Catriona MacCallum), but other than providing general advice during the design of the study and organizing the provision of PDFs of included manuscripts, they had no role.

Funding/Support: The Medical Research Council, National Centre for the Replacement Refinement and Reduction of Animals in Research, Biotechnology and Biological Sciences Research Council, and Wellcome Trust pooled resources without a normal grant cycle to fund this project.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the abstract. The funders used their social media streams to publicize the study and recruit outcome assessors. National Centre for the Replacement, Refinement, and Reduction of Animals in Research employees were not allowed to enroll as outcome assessors because of their possible conflict of interest as sponsors of the ARRIVE guidelines.

Group Members: The IICARus Collaborative group includes the following members: *University of Edinburgh, Edinburgh, UK:* Emily Sena, Cadi Irvine, Kaitlyn Hair, Fala Cramond, Paula Grill, Gillian Currie, Alexandra Bannach-Brown, Zsanett Bahor, Daniel-Cosmin Marcu, Monica Dingwall, Victoria Hohendorf, Klara Zsofia Gerlei, Victor Jones, Anthony Shek, David Henshall, Emily Wheeler, Edward Christopher, and Malcolm Macleod; *University of Tasmania, Hobart, Tasmania:* David Howells; *University of Nottingham, Nottingham, UK:* Ian Devonshire and Philip Bath; *Public Library of Science, Cambridge, UK:* Catriona MacCallum; *Imperial College London, London, UK:* Rosie Moreland; *Mansoura University, Mansoura, Egypt:* Sarah Antar, Mona Hosh, and Ahmed Nazzal; *University of New South Wales, Kensington, NSW, Australia:* Katrina Blazek; *Animal Sciences Unit, Animal and Plant Health Agency, Addlestone, UK:* Timm Konold; *University of Glasgow, Glasgow, UK:* Terry Quinn and Teja Gregorc; *AstraZeneca, Wilmington, Delaware, USA:* Natasha Karp; *Nuffield Research Placement Student, London, UK:* Privjyot Jheeta and Ryan Cheyne; *GlaxoSmithKline, Middlesex, UK:* Joanne Storey; *University College London, London, UK, and École Normale Supérieure, Paris, France:* Julija Baginskaite; *University Medical Center Utrecht, Utrecht, the Netherlands:* Kamil Laban; *University of Rome Sapienza, Rome, Italy:* Arianna Rinaldi; *Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands:* Kimberley Wever; *University of Southampton, Southampton, UK:* Savannah Lynn; *Federal University of Rio de Janeiro, Rio de Janeiro, Brazil:* Evandro Araújo De-Souza; *University of Birmingham, Birmingham, UK:* Leigh O'Connor; *Hospital Research Center of the Sacred Heart of Montreal, Montreal, QC, Canada:* Emmanuel Charbonney; *National Cancer Institute, Milano, Italy:* Marco Cascella; *Federal University of Santa Catarina, Florianópolis, Brazil:* Cilene Lino de Oliveira; *University of Geneva, Geneva, Switzerland:* Zeinab Ammar; *British American Tobacco, London, UK:* Sarah Corke; *Ministry of Health, Cairo, Egypt:* Mahmoud Warda; *Vita-Salute San Raffaele University, Milan, Italy:* Paolo Roncon; *University of Hertfordshire, Hertfordshire, UK:* Daniel Baker; *University of Veterinary Medicine Hanover, Hanover, Germany:* Jennifer Freymann.

Trial Registration

Association of Trial Registration With Reporting of Clinical Trials: Comparison of Protocols, Registries, and Published Articles

An-Wen Chan,^{1,2} Annukka Pello,³ Jessica Kitchen,¹ Anna Axentiev,¹ Jorma Virtanen,⁴ Annie Liu,¹ Elina Hemminki⁵

Objective To evaluate adherence to trial registration and its association with subsequent publication and selective reporting of primary outcomes in an unselected cohort of clinical trials.

Design This was an inception cohort study of all initiated clinical trial protocols approved in 2002 (n=135) and 2007 (n=113) by the research ethics committee for the region of Helsinki and Uusimaa, Finland. We identified registry records and articles published up to February 2017 using keywords to search trial registries, PubMed, EMBASE, and Google. Trial characteristics (approval year, funding, sample size, intervention type, number of arms and centers) and outcomes were abstracted from each protocol, registry record, and publication. Using descriptive statistics and multivariable logistic regression, we determined the rates and predictors of registration and publication; the proportion of trials with discrepant primary outcomes in the protocol compared with the registry and publication; and the association between registration and subsequent publication without discrepant primary outcomes. Discrepancies were defined as (1) a new primary outcome being reported that was not specified as primary in the protocol; or (2) a protocol-defined primary outcome being omitted or downgraded (reported as secondary or unspecified) in the registry or published article.

Results Registration rates increased from 0% (0 of 135) for trials approved in 2002 to 61% (69 of 113) in 2007. Overall, 130 of 248 of all trials (52%) were published (publication years 2003 through 2016); 16 of 69 registered trials (23%) had discrepancies in primary outcomes defined in the registry compared with the protocol, while 24 of 116 published trials (21%) had discrepancies in primary outcomes between the published article and the protocol. Among trials approved in 2007, trial registration was significantly associated with subsequent publication (68% of registered trials vs 39% of unregistered trials; adjusted odds ratio [aOR], 4.5; 95% CI, 1.1-18). Registered trials were also significantly more likely than unregistered trials to be subsequently published with the same primary outcomes defined in the published article compared with the protocol (64% vs 25%; aOR, 5.8; 95% CI, 1.4-24).

Conclusions Clinical trials are not only often unregistered and unpublished but also discrepant in the reporting of primary outcomes across different information sources. These major deficiencies impair transparency and facilitate the biased reporting of trial results, which can be mitigated through adherence to trial registration. Journal editors, legislators, funding agencies, regulators, research ethics committees, and sponsors should implement and enforce

policies mandating registration and public access to full protocols for all clinical trials.

¹Women's College Research Institute, Women's College Hospital, Toronto, Ontario, Canada, anwen.chan@utoronto.ca; ²Department of Medicine, University of Toronto, Ontario, Canada; ³Faculty of Medicine, University of Helsinki, Helsinki, Finland; ⁴Faculty of Medicine, University of Oulu, Oulu, Finland; ⁵THL (National Institute for Health and Welfare), Helsinki, Finland

Conflict of Interest Disclosures: None reported.

Funding/Support: This project was supported by the Canadian Institutes of Health Research Dissemination Events (grants MET 117434 and MET 133851) and the Academy of Finland (grant No. 28356).

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the abstract.

Impact of FDAAA on Registration, Results Reporting, and Publication of Neuropsychiatric Clinical Trials Supporting FDA New Drug Approval, 2005-2014

Constance X. Zou,¹ Jessica E. Becker,^{2,3,4} Adam T. Phillips,⁵ Harlan M. Krumholz,^{6,7,8,9} Jennifer E. Miller,¹⁰ Joseph S. Ross^{7,8,9}

Objective Selective publication and reporting of clinical trial results undermines evidence-based medicine. The 2007 Food and Drug Administration Amendments Act (FDAAA) mandates, with few exceptions, the registration and reporting of results of all non-phase I clinical trials on ClinicalTrials.gov for approved products. The objective of this study was to determine whether efficacy trials supporting US Food and Drug Administration (FDA) approval of new drugs used for neurological and psychiatric conditions that were completed after FDAAA was enacted were more likely to have been registered, have their results reported, and be published in journals than those completed pre-FDAAA.

Design We conducted a retrospective observational study of efficacy trials reviewed by the FDA as part of any new neuropsychiatric drugs approved between 2005 and 2014. In January 2017, for each trial, we searched ClinicalTrials.gov for the registration record and for reported results, and we searched MEDLINE-indexed journals using PubMed for corresponding publications. In addition, published findings were validated against FDA interpretations described in regulatory medical review documents. Trials were considered FDAAA applicable if they were initiated after September 27, 2007, or were still ongoing as of December 26, 2007. The rates of trial registration, results reporting, publication, and publication-FDA agreement were compared between pre-FDAAA and post-FDAAA trials using Fisher exact test.

Results Between 2005 and 2014, the FDA approved 37 new neuropsychiatric drugs on the basis of 142 efficacy trials, of which 41 were FDAAA applicable. Post-FDAAA trials were significantly more likely to be registered (100% vs 64%; $P < .001$) and to report results (100% vs 10%; $P < .001$) than

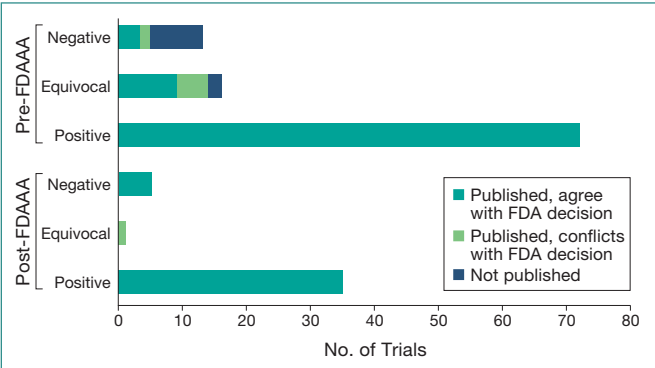
pre-FDAAA trials, but post-FDAAA trials were not significantly more likely to have been published (100% vs 90%; $P = .06$) nor to have been published with findings in agreement with the FDA's interpretation (98% vs 93%; $P = .28$) (**Figure 4**). Subgroup analyses suggest that the changes in overall publication rate were primarily the consequence of publishing negative trials, as all pre-FDAAA and post-FDAAA positive trials were published (72 of 72 and 35 of 35, respectively), whereas 38% (5 of 13) of pre-FDAAA negative trials were published vs 100% (5 of 5) of post-FDAAA negative trials.

Conclusions After FDAAA was enacted, all efficacy trials reviewed by the FDA as part of new drug applications for neuropsychiatric drugs were registered, with the results reported and published. Moreover, nearly all were published with interpretations that agreed with the FDA's interpretation. While our study was limited by searching for registration status only on ClinicalTrials.gov, our findings suggest that by mitigating selective publication and reporting of clinical trial results, FDAAA improved the availability of evidence for physicians and patients to make informed decisions regarding the care of neuropsychiatric illnesses.

¹Yale School of Medicine, New Haven, CT, USA, constance.zou@yale.edu; ²Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA; ³McLean Hospital, Belmont, MA, USA; ⁴Harvard Medical School, Boston, MA, USA; ⁵Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA; ⁶Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA; ⁷Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, New Haven, CT, USA; ⁸Robert Wood Johnson Foundation Clinical Scholars Program, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA; ⁹Department of Health Policy and Management, Yale School of Public Health, New Haven, CT, USA; ¹⁰Division of Medical Ethics, Department of Population Health, NYU School of Medicine, Bioethics International, New York, NY, USA

Conflict of Interest Disclosures: Constance Zou has received a fellowship through the Yale School of Medicine from the National Heart, Lung, and Blood Institute. Jennifer Miller has received research support through New York University from the Laura and John Arnold Foundation to support the Good Pharma Scorecard. Harlan Krumholz and Joseph Ross have received research support

Figure 4. Pre- and Post-FDAAA Efficacy Trials Supporting Neuropsychiatric Drugs First Approved Between 2005 and 2014: Publication Status and Published Conclusion Concordance With FDA Decision



FDA indicates the US Food and Drug Administration; FDAAA, the 2007 Food and Drug Administration Amendments Act.

through Yale University from Johnson and Johnson to develop methods of clinical trial data sharing, from Medtronic and the US Food and Drug Administration (FDA) to develop methods for postmarket surveillance of medical devices, and from the US Centers of Medicare and Medicaid Services to develop and maintain performance measures that are used for public reporting. Harlan Krumholz also has received compensation as a member of the Scientific Advisory Board for United Healthcare. Joseph Ross has received research support through Yale University from the FDA to establish a Center for Excellence in Regulatory Science and Innovation at Yale University and the Mayo Clinic, from the Blue Cross Blue Shield Association to better understand medical technology evaluation, and from the Laura and John Arnold Foundation to support the Collaboration on Research Integrity and Transparency at Yale University.

Evaluation of the ClinicalTrials.gov Results Database and Its Relationship to the Peer-Reviewed Literature

Deborah A. Zarin,¹ Tony Tse,¹ Rebecca J. Williams,¹ Thiyaagu Rajakannan,¹ Kevin M. Fain¹

Objective As of February 22, 2017, ClinicalTrials.gov contained summary results for 24,377 studies and received 160 new submissions weekly. We estimate that US academic medical centers are required to report more than half of their sponsored trials to ClinicalTrials.gov under federal policies. We previously estimated that one-half of registered studies with results posted on ClinicalTrials.gov lacked results publications. It is critical to continue assessing the degree to which this database meets its intended goals. The objective of this study was to assess the potential scientific impact of the ClinicalTrials.gov results database using our 2013 evaluation framework.

Design We analyzed 2 samples of ClinicalTrials.gov results data to assess the impact on the available evidence base.

Results On February 10, 2017, 10,464 of 24,251 posted results (43%) had links to PubMed. Because not all publications are automatically linked and not all linked publications report results, we manually examined a random sample of 100 sets of posted results listing study completion dates in 2014. Of these, 28 had at least 1 results publication prior to results posting, 15 had a results publication after results posting, and we could not identify results publications for 57 studies. We also identified examples of how publications leveraged the information on ClinicalTrials.gov. To further examine the potential impact on selective publication, we evaluated drug-condition-sponsor “families.” We identified 329 registered, industry-funded, phase 2 through 4, US trials completed or terminated from 2007 through 2009, representing 88 drugs and 96 unique drug-condition-sponsor families (eg, Amgen-sponsored trials of alendronate for osteoporosis). Ideally, summary results for all trials in all families would be publicly available. As of December 1, 2014, of 329 trials, 109 (33%) had results posted on ClinicalTrials.gov only, 42 (13%) available from PubMed only, 81 (25%) available from both, and 97 (29%) in neither (**Table 12**). Overall, 45 of the 96 drug-condition-sponsor families had results available for all 144 trials from at least 1

Table 12. Study Design Characteristics of 329 ClinicalTrials.gov-Registered Trials in the Drug-Condition-Sponsor “Families” Sample by Results Dissemination Category as of April 27, 2017

Study Design Characteristic	Trials by Dissemination of Results, No. (% by Characteristic)				
	Results Disclosure			Any vs No Results Disclosure on ClinicalTrials.gov and PubMed	
	ClinicalTrials.gov Only (n = 109)	PubMed Only (n = 42)	Both ClinicalTrials.gov and PubMed (n = 81)	Results Disclosure Total (n = 232)	Neither ClinicalTrials.gov nor PubMed (n = 97)
Interventional model					
Parallel assignment (n = 242)	79 (33)	33 (14)	68 (25)	180 (74)	62 (26)
Single group assignment (n = 72)	26 (36)	7 (10)	8 (11)	41 (57)	31 (43)
Crossover assignment (n = 9)	2 (22)	1 (11)	4 (44)	7 (78)	2 (22)
Factorial assignment (n = 4)	2 (50)	0	1 (25)	3 (75)	1 (25)
Missing data (n = 1)	0	1 (50)	0	0 (50)	1 (50)
Masking					
Open label (n = 99)	37 (37)	9 (9)	18 (18)	64 (65)	35 (35)
Double blind (n = 209)	62 (60)	32 (15)	57 (27)	151 (72)	58 (28)
Single blind (n = 11)	6 (55)	0	2 (18)	8 (73)	3 (27)
Missing data (n = 10)	4 (40)	1 (10)	4 (40)	9 (90)	1 (10)
Allocation					
Randomized (n = 246)	75 (30)	34 (14)	70 (28)	179 (73)	67 (27)
Missing data (n = 30)	17 (57)	4 (13)	1 (3)	22 (73)	8 (27)
Nonrandomized (n = 53)	17 (32)	4 (8)	10 (19)	31 (58)	22 (42)
No. of sites					
Multiple (n = 218)	70 (32)	29 (13)	62 (28)	161 (74)	57 (26)
Single (n = 58)	19 (33)	6 (10)	3 (5)	28 (48)	30 (52)
Missing data (n = 53)	20 (38)	7 (13)	16 (30)	43 (81)	10 (19)
No. of participants enrolled					
1-100 (n = 102)	37 (36)	13 (13)	11 (11)	61 (60)	41 (40)
101-500 (n = 150)	52 (35)	21 (14)	32 (21)	105 (70)	45 (30)
>500 (n = 75)	20 (27)	8 (11)	38 (51)	66 (88)	9 (12)
Missing data (n = 2)	0	0	0	0	2 (100)

source, 18 families involving a total of 48 trials had no results available, and 15 families had results disclosed on ClinicalTrials.gov only.

Conclusions Between 33% (109 of 329) and 57% (57 of 100) of completed or terminated ClinicalTrials.gov-registered trials have posted results but no corresponding PubMed-cited results articles. These findings suggest that ClinicalTrials.gov provides a unique source of results for substantial numbers of trials.

¹ClinicalTrials.gov, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA, dzarin@nih.gov

Conflict of Interest Disclosures: All authors work for ClinicalTrials.gov as full-time employees of the National Library of Medicine. Dr Zarin is a member of the Peer Review Congress Advisory Board but was not involved in the review or decision for this abstract.

Funding/Support: This work was supported by the Intramural Research Program of the National Library of Medicine, National Institutes of Health. The National Library of Medicine has approved this submission.

Funding/Grant Review

Geographic and Gender Bias in Peer Review of Applications Submitted to the Swiss National Science Foundation

João Martins,¹ François Delavy,¹ Anne Jorstad,¹ Matthias Egger¹

Objective The Swiss National Science Foundation (SNSF), the leading public research funder in Switzerland, relies on external experts to review grant applications. Applicants can propose reviewers, provided there are no obvious conflicts of interests. On average, applications receive 3 reviews, 1 of which is typically from a reviewer proposed by the applicants. We examined whether the source of the review, the gender of the principle applicant and the reviewer, and the country of affiliation of reviewers influenced the scores given to grant applications submitted to the SNSF.

Design Reviewers scored applications from 1 (poor) to 6 (outstanding). We calculated mean scores by source of

reviewers (applicant vs SNSF), country of affiliation of reviewers (Switzerland vs international), and gender of applicants and reviewers. We fit a multivariable linear regression model adjusting for all these variables plus calendar year of submission, discipline (21 disciplines), and applicants' age (5 age classes) and affiliation (4 institution types).

Results Between 2009 and 2015, 36,993 reviewers assessed 12,132 applications for the SNSF. The mean (SD) score of reviewers proposed by applicants (n=8308) was 5.12 (1.01) vs 4.47 (1.25) for reviewers proposed by the SNSF (n=26,594). Mean (SD) scores were 4.19 (1.27) for Swiss experts (n=8399) vs 4.76 (1.19) for international experts (n=26,503); 4.44 (1.25) for female (n=7121) vs 4.67 (1.22) for male (n=27,781) principle applicants; and 4.48 (1.26) for reviews from female (n=6933) vs 4.66 (1.22) from male (n=27,969) reviewers. In adjusted analyses, the gender differences were attenuated, whereas the other differences changed little (**Table 13**). All differences were statistically significant.

Conclusions Applications received higher scores from applicant-proposed reviewers and lower scores from Swiss-based experts. Scores were lower for applications submitted by female applicants. Our results are compatible with a positive bias of reviewers chosen by the applicant, or a negative bias of experts based in Switzerland, and cannot exclude bias against female applicants. Interestingly, female reviewers consistently scored applications lower than male reviewers, independent of the applicant's gender. Panels making funding decisions should be aware of these potential biases. Given the association between scores and source of reviewer, the SNSF no longer accepts reviewers proposed by the applicants.

¹Swiss National Science Foundation, Bern, Switzerland, joao.martins@snf.ch

Conflict of Interest Disclosures: The authors are employees of the Swiss National Science Foundation.

Table 13. Unadjusted and Adjusted Differences in Scores Assigned by Reviewers of Grant Applications Submitted to the Swiss National Science Foundation

Variable	Difference (95% CI) ^a	
	Unadjusted	Adjusted
Source of reviewer		
Applicant vs SNSF	0.65 (0.62 to 0.68)	0.52 (0.49 to 0.55)
Affiliation of reviewer		
Switzerland vs international	−0.56 (−0.59 to −0.53)	−0.50 (−0.53 to −0.47)
Gender of applicant		
Female vs male	−0.23 (−0.19 to −0.26)	−0.09 (−0.06 to −0.12)
Gender of reviewer		
Female vs male	−0.17 (−0.13 to −0.21)	−0.07 (−0.03 to −0.10)

^aAll unadjusted *P* values from *t* tests <.001. Adjusted results from a linear regression model adjusted for calendar year of submission, discipline, and applicants' age and affiliation; all adjusted *P* values <.001.

Stakeholder Perceptions of Peer Review at the National Institutes of Health Center for Scientific Review

Mary Ann Guadagno,¹ Richard K. Nakamura¹

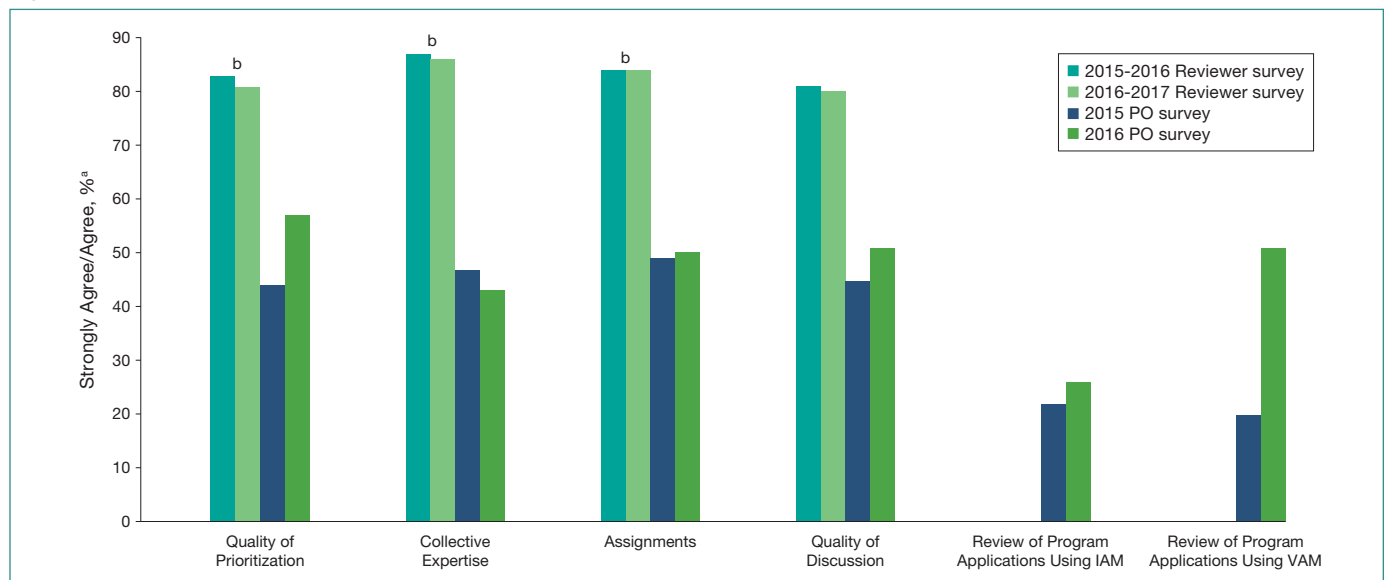
Objective To identify best practices for the successful peer review of grant applications and areas for improvement at the National Institutes of Health (NIH) Center for Scientific Review (CSR), the following questions guided an evaluation study: (1) to what extent are current CSR practices for peer review optimal for achieving its mission? and (2) what are the areas of success and improvement in the quality of peer review?

Design Pilot assessments were conducted to develop a short “Quick Feedback” survey instrument with four 7-point Likert-type scale statements ranging from “strongly agree” to “strongly disagree,” measuring key features of peer review, and an open text box for comments. During 1 grant cycle between 2015–2016 and 2016–2017, 2 surveys were sent to 10,262 and 10,228 reviewers, respectively, in all CSR study sections. In 2015, a survey was sent to 916 NIH Program Officers (POs), and a replication survey was sent to POs in 2016 to 905 POs. During 2015, 27 focus groups were conducted with 4 stakeholder groups, and 10 personal interviews were completed with NIH Institute Directors. Focus group participants were selected from NIH databases to ensure diversity. Interrater reliability between coders was 95.8%.

Results The 2015–2016 reviewer survey yielded a response rate of 47.1% (4832 of 10,262), and the 2016–2017 reviewer survey yielded a response rate of 47.0% (4807 of 10,228). The 2015 PO survey had a response rate of 38.0% (348 of 916), and the 2016 replication PO survey yielded a response rate of 37.0% (335 of 905). Nonrespondents were not substantially different from respondents. “Quick Feedback” surveys with reviewers in both years reported a high level of satisfaction with the peer review process. More than 80% of reviewers indicated they either “strongly agreed” or “agreed” that panels were doing a good job in terms of scoring and discussion and CSR did a good job relative to the quality of the rosters and assignments (**Figure 5**). Program Officers were less favorable than reviewers in both years, with only 43% to 57% of POs responding favorably. Program Officers' dissatisfaction with review meetings focused on insufficient reviewer expertise in general and technical and logistical challenges at meetings more specifically. Focus group results supported these findings. Areas for improvement included reducing the burden of peer review for all stakeholders, technical and logistical issues during meetings, need for clearer communication, and more guidance on preparing applications.

Conclusions A comprehensive evaluation using systematic surveys, focus groups, and interviews has resulted in useful suggestions for improving best practices for peer review by stakeholders in real time. Areas of success and suggestions for

Figure 5. Overall Center for Scientific Review Quick Feedback Favorable Responses



The 2015-2016 reviewer survey yielded a response rate of 47.1% (4832 of 10,262), and the 2016-2017 reviewer survey yielded a response rate of 47.0% (4807 of 10,228). The 2015 Program Officer (PO) survey had a response rate of 38.0% (348 of 916), and the 2016 replication PO survey had a response rate of 37.0% (335 of 905). IAM indicates Internet-assisted meeting; VAM, video-assisted meeting.

*Strongly agree or agree refers to a 1 or 2, respectively, as assessed on a 7-point Likert-type scale.

^bIAM reviewers not included in 2016.

improvements by stakeholders are being addressed by leadership.

¹Center for Scientific Review, National Institutes of Health, Bethesda, MD, USA, guadagma@csr.nih.gov

Conflict of Interest Disclosures: Both authors are federal employees at the National Institutes of Health. Survey research was conducted as part of their federal employment responsibilities.

Funding/Support: Focus groups and personal interviews were funded by the NIH Evaluation Set-Aside Program (14-5725 CSR), administered by the Office of Program Evaluation and Performance of the National Institutes of Health.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the abstract.

Testing of 2 Application Ranking Approaches at the National Institutes of Health Center for Scientific Review

Richard K. Nakamura,¹ Amy L. Rubinstein,¹ Adrian P. Vanea,¹ Mary Ann Guadagno¹

Objective The National Institutes of Health (NIH) is a US agency that distributes approximately \$20 billion each year for research awards based on a rigorous peer review that provides a merit score for each application. Final scores are based on the mean of scores from reviewers and then ranked via percentile. In 2009, the NIH changed its scoring system from a 40-point scale to a 9-point scale. There have been concerns that this new scale, which is functionally cut in half for the 50% of applications that are considered competitive, is not sufficient to express a study section's judgment of relative merit. The question guiding these pilot studies was whether

alternative methods of prioritizing applications could reduce the number of tied scores or increase ranking dispersal.

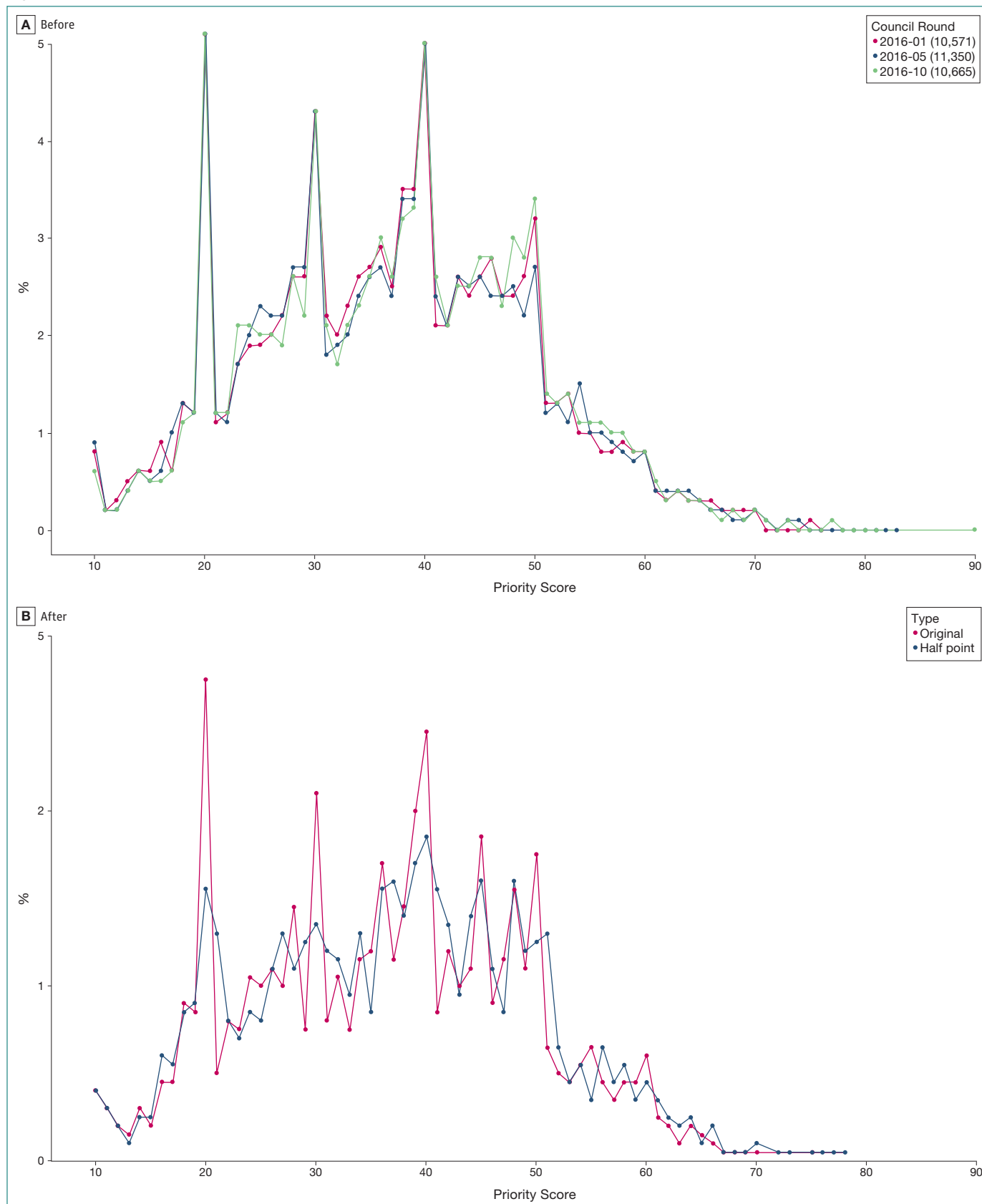
Design The Center for Scientific Review has been testing alternate scoring systems, including (A) postmeeting ranking of the top scoring applications, in which reviewers rank-order the 10 best applications at the end of a review, and (B) giving reviewers the option of adding or subtracting a half point during final scoring of applications following discussion. These alternatives were compared against standard scoring in real study sections to see if they could improve prioritization of applications and reduce the number of tied scores. Reviewer opinions of the ranking systems were assessed, including surveys for alternate B.

Results (A) Postmeeting ranking of applications were applied to 836 applications across 32 study sections; these often produced rankings inconsistent with scores given to applications by reviewers. The best 2 or 3 scored applications were generally agreed on, but increased disagreement among reviewers was observed with poorer average scores. Most reviewers liked the ranking system, but there was more hesitation about recommending adoption of this practice over the current scoring system. (B) Making available a half point to add or subtract freely in final voting was applied to 1371 applications across 39 study sections; the half-point system helped to spread scores and halved the number of ties (**Figure 6**). It was also recommended for adoption by 72% of reviewers in postmeeting surveys.

Conclusions Initial results of the half-point scoring system have been interpreted favorably. The Center for Scientific Review will conduct a full test of the half-point scoring system under real review conditions.

¹Center for Scientific Review, National Institutes of Health, Bethesda, MD, USA, rnakamur@mail.nih.gov

Figure 6. Distribution of National Institutes of Health Grant Application Scores by Percent Before and After Use of the Half-Point Option



A, Distribution of final scores for grant applications as a percent of all scores (of 32,586 applications). Each application received scores from many reviewers that were multiplied by 10 and averaged to the nearest unit. Possible final scores for each application ranged from 10 to 90. Dates refer to the cycle of review and the number is the quantity of applications with scores. In January 2016, 10,571 applications received scores; in May 2016, 11,350 applications received scores; and in October 2016, 10,665 applications received scores. B, Comparison of distribution of original average scores (red) with scores for which reviewers were allowed to add or subtract a half point (blue). Average scores are rounded to the nearest digit to establish ranking. Original refers to the proportion of scores at each possible score level under normal whole digit scoring. Half point refers to the proportion of scores at each possible score level when reviewers used whole digits plus or minus 1 half point.

Conflict of Interest Disclosures: All authors are employees of the National Institutes of Health.

Funding/Support: Study funded by the National Institutes of Health.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the abstract.

Scientist, Patient, and Stakeholder Roles in Research Application Review: Analysis of the Patient-Centered Outcomes Research Institute (PCORI) Approach to Research Funding

Laura P. Forsythe,¹ Lori B. Frank,¹ A. Tsahai Tafari,¹ Sarah S. Cohen,² Michael Lauer,^{1,3} Steve Clauser,¹ Christine Goertz,^{1,4} and Suzanne Schrandt¹

Objective Scientific review of funding applications was established to fund rigorous, high-impact research. The Patient-Centered Outcomes Research Institute (PCORI) uses unique review criteria and includes patients and other healthcare stakeholders as reviewers. This study assesses the relative importance of each criterion and the associations of different reviewer types' ratings with final scores and funding outcomes.

Design This study is a retrospective, cross-sectional analysis of PCORI Merit Review administrative data for 5 funding cycles from 2013 through 2015. Before a panel discussion, patients and other stakeholders were required to score each application overall and on 3 criteria: potential to improve care and outcomes, patient-centeredness, and engagement. Scientist reviewers also scored impact of condition and technical merit. Scores ranged from 1 (exceptional) to 9 (poor). All reviewers provided postdiscussion overall scores. Funding decisions were made by the PCORI Board of Governors based on Merit Review, portfolio balance, and programmatic fit. Linear regression models stratified by reviewer type (ie, scientist, patient, or other stakeholder) tested associations of postdiscussion overall scores with prediscussion criteria scores. Associations between funding decisions and prediscussion criteria scores were tested using logistic regression. All models adjusted for funding program, review cycle, and principal investigator characteristics (ie, National Institutes of Health funding, clinical degree[s] of applicants, and years of experience of applicants).

Results A total of 535 reviewers (254 scientists, 139 patients, and 142 stakeholders) reviewed 1312 applications; 663 (50.5%) were discussed and 121 (9.2%) were funded. Prediscussion mean (SD) overall scores were higher (ie, worse) for scientist reviewers (4.9 [2.1]) than patient reviewers (4.2 [2.2]) and stakeholder reviewers (4.2 [2.1]) ($P < .001$). The mean overall score postdiscussion was 28.0 for funded applications and 50.1 for unfunded applications. All reviewer types changed their overall score through panel discussion for more than half of the applications. Score agreement across reviewer types was greater postdiscussion. For all reviewer types, postdiscussion review scores were

positively associated with at least 1 prediscussion criterion score from each of the 3 reviewer types (**Table 14**). The strongest association with postdiscussion overall scores for all reviewer types was scientists' ratings of technical merit. More favorable prediscussion ratings by each reviewer type for the potential to improve care and outcomes and scientist reviewers' ratings of technical merit and patient-centeredness were associated with greater likelihood of funding.

Conclusions Scientist, patient, and stakeholder views of applications converged following discussion. Technical merit is critical to funding success, but patient and stakeholder ratings of other criteria also relate to application disposition. Results suggest that research application review can incorporate nonscientist perspectives in scoring and funding outcomes.

¹Patient-Centered Outcomes Research Institute, Washington, DC, USA, lforsythe@pcori.org; ²EpidStat Institute, Ann Arbor, MI, USA; ³National Institutes of Health, Bethesda, MD, USA; ⁴Palmer Center for Chiropractic Research, Davenport, IA, USA

Conflict of Interest Disclosures: None reported.

Funding/Support: This work was funded by the Patient-Centered Outcomes Research Institute (PCORI).

Role of the Funder/Sponsor: Members of the PCORI staff, Board of Governors, and Methodology Committee designed and conducted the study and reported the results.

Tuesday, September 12, 2017

Innovations in Peer Review and Scientific Publication

Peer Review Innovations

Assessment of Author Demand for Double-blind Peer Review in IOP (Institute of Physics) Publishing Journals

Simon Harris,¹ Marc Gillett,¹ Pernille Hammelsoe,¹ and Tim Smith¹

Objective The main objective of this study is to generate market intelligence on author demand for double-blind peer review (DBPR) in areas of study where single-blind review is the norm, namely, materials science and biomedical physics and/or engineering. In addition to assessing authors' perception of the double-blind model and their satisfaction with the process, the pilot study will also collect data to compare operational aspects of the peer-review process between the single-blind and double-blind models.

Design Alternative submission and peer review sites have been set up for the journals Materials Research Express (MRX) and Biomedical Physics & Engineering Express (BPEX). Authors are able to choose between submitting an article for double-blind or single-blind review. Authors are responsible for anonymizing their manuscript before submitting it for DBPR. The pilot scheme runs for 1 year, from January 2017 to December 2017. A full analysis of data

Table 14. Multiple Regression Analysis for Postdiscussion Overall Scores and Funding Decision in Relation to Individual Reviewer Criteria (n = 635 Applications)^a

Characteristic	Linear Regression Models: Postdiscussion Overall Score Stratified by Reviewer Type ^b						Logistic Regression Model for Funding (Yes vs No)	
	Model 1: Scientist Reviewers' Overall Scores (<i>R</i> ² = 0.58)		Model 2: Patient Reviewers' Overall Scores (<i>R</i> ² = 0.47)		Model 3: Stakeholder Reviewers' Overall Scores (<i>R</i> ² = 0.51)			
	Estimate (SE)	<i>P</i> Value	Estimate (SE)	<i>P</i> Value	Estimate (SE)	<i>P</i> Value	Odds Ratio (95% CI)	<i>P</i> Value
Scientist prediscussion scores for review criteria ^c								
Impact	0.03 (0.04)	.46	0.04 (0.05)	.36	−0.04 (0.05)	.41	1.09 (0.76-1.56)	.64
Potential to improve care and outcomes	0.30 (0.04)	<.001	0.21 (0.05)	<.001	0.29 (0.05)	<.001	0.42 (0.29-0.59)	<.001
Technical merit	0.41 (0.04)	<.001	0.29 (0.04)	<.001	0.37 (0.04)	<.001	0.45 (0.34-0.60)	<.001
Patient-centeredness	0.10 (0.04)	.02	0.08 (0.05)	.09	0.11 (0.05)	.02	0.61 (0.41-0.89)	.01
Engagement	0.10 (0.04)	.009	0.07 (0.03)	.12	0.05 (0.04)	.29	0.95 (0.69-1.28)	.68
Patient prediscussion scores for review criteria ^c								
Potential to improve care and outcomes	0.12 (0.03)	<.001	0.15 (0.03)	<.001	0.08 (0.03)	.01	0.77 (0.64-0.94)	.01
Patient-centeredness	0.05 (0.03)	.09	0.05 (0.03)	.15	0.05 (0.03)	.09	0.85 (0.70-1.04)	.12
Engagement	0.02 (0.03)	.44	0.07 (0.03)	.02	0.04 (0.03)	.17	0.96 (0.81-1.14)	.65
Stakeholder prediscussion scores for review criteria ^c								
Potential to improve care and outcomes	0.08 (0.03)	.007	0.15 (0.03)	<.001	0.12 (0.03)	<.001	0.75 (0.62-0.92)	.004
Patient-centeredness	0.05 (0.03)	.09	0.08 (0.03)	.02	0.09 (0.03)	.005	0.81 (0.64-1.01)	.06
Engagement	0.08 (0.03)	.005	0.11 (0.03)	<.001	0.10 (0.03)	.001	0.89 (0.72-1.09)	.26

^aAll models adjusted for program, review cycle, and characteristics of the principal investigator, including any/none National Institutes of Health funding, any/none clinical degree(s), and years of experience since attaining terminal degree (1–4, 5–9, or 10+). Mean scores were used for the 2 scientist reviewers for this analysis. Applications scored on a scale from 1 (exceptional) to 9 (poor).

^bOverall score from all reviewers of a given type for the entire in-person panel.

^cReview criteria scores from preliminary reviewers only.

will take place after 6 months, and the final analysis will take place after 12 months. The percentage of the total of direct submissions to each journal that are submitted for DBPR will be measured and reported. This uptake figure is the primary measure of this study, and prior to the study a significant uptake threshold was set at 10% for MRX and 20% for BPEX. Surveys to submitting authors at the final decision stage measure satisfaction levels with the DBPR process and will be compared with single-blind author satisfaction levels. Double-blind authors are also asked why they chose DBPR rather than single-blind peer review.

Results Initial data from January 2017 to May 2017 show that 20% of direct submissions to MRX were double blind (137 of 677 direct submissions), and BPEX shows a similar uptake (9 of 46 submissions). Peer-review times are slightly shorter for DBPR than for SBPR. Eight authors who chose DBPR have responded to the survey, and 7 of them stated that they chose DBPR because they think it is the most fair. The average satisfaction level with the DBPR experience is 8.5 out of a 10-point scale.

Conclusions These initial results suggest that there is a significant demand from authors for DBPR in these communities. If anything, we feel that this study will underestimate the demand given that DBPR is voluntary and most authors submitting to these journals are not normally in the habit of anonymizing their manuscripts. Further

conclusions will be drawn based on the 6-month results in July 2017, in time for the Peer Review Congress in September.

¹IOP Publishing, Bristol, UK, simon.harris@iop.org

Conflict of Interest Disclosures: All authors are full-time employees of IOP Publishing. No other conflicts were reported.

Use of Open Review by Discipline, Country, and Over Time: An Analysis of Reviews and Journal Policies Posted on Publons

Sarah Parks,¹ Salil Gunashekar,¹ Elta Smith¹

Objective To understand what forms of open peer review are being used and how these vary by discipline, by country, and over time.

Design We used data from Publons to explore our objectives. Publons is a publisher neutral platform that allows users to record peer reviewer activity and their preferences for signing and/or publishing reviews and also records journal policies on peer review. Publons publishes reviews openly if the reviewer has selected this option, provided the journal policy permits this. We focused on the 7458 journals with at least 10 reviews recorded in Publons as of November 2016. Publons also contains information on the peer review policies of approximately 12% of journals (3692 of 30,000 assuming there are 30,000 English-language journals in the world). We split the data by the country of the reviewer (where there

are at least 100 distinct reviewers) and the discipline of the journal to look at trends in use of open peer review, including the difference between the wishes of reviewers to publish their reviews and journal policies and the current state of policies for open peer review for journals in different fields. The disciplines assessed included life sciences, natural sciences, engineering sciences, humanities, and social and behavioral sciences.

Results Of the 3692 journals analyzed, 3.5% (n=130) allow reviewers to sign their reviews, and 2.3% (n=85) allow or require peer reviews to be published. There are 474,036 reviews in Publons from the 7458 journals that have at least 10 reviews recorded in Publons. Of these reviews, 1.7% (n=7857) are published openly. For an additional 7904 reviews (1.7% of total reviews), the reviewers have indicated that they would prefer Publons to publish the review openly, but the journal policy does not permit this. The rate of published open reviews in Publons varies by country, ranging from 0% in Argentina (0 of 1,076), Czech Republic (0 of 867), Ireland (0 of 1667), and Romania (0 of 1392) to 14.1% (149 of 1058) in Saudi Arabia, and by discipline, ranging from 0.8% (205 of 26,762) for engineering sciences to 2.9% (18 of 615) for humanities.

Conclusions Open peer review is happening in different forms on the Publons platform, but it still accounts for a small percentage of reviews carried out. While these data provide insight into the use of open peer review, it is important to note that the data are sourced from Publons and may not be representative of all scholarly disciplines, countries, open reviews, and open reviewer activity. In addition, we were not able to differentiate the posting of full reviews vs the posting of metadata of reviewer activity. Future research could address these limitations and other aspects of open review.

¹RAND Europe, Cambridge, UK, sparks@rand.org

Conflict of Interest Disclosures: None reported.

Funding/Support: Initial work on this topic was funded by the European Commission as part of a project to develop a monitor for open science.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the abstract.

Acknowledgments: We would like to thank Publons Limited and Publons UK Limited for providing us with data.

Comparison of Acceptance of Peer Reviewer Invitations by Peer Review Model: Open, Single-blind, and Double-blind Peer Review

Maria Kowalczyk,¹ Michelle Samarasinghe²

Objective Anecdotal evidence from editors suggests it is more difficult to recruit reviewers for journals that use fully open peer review compared with single- or double-blind peer review. The aim of this study was to determine whether there is a difference in the proportion of reviewers who agree to undertake peer review of manuscripts for journals that use different peer review models in different subject areas.

Design Retrospective analysis of BioMed Central and SpringerOpen journals that use the 3 different peer review models in biomedicine, chemistry, clinical medicine, computer science, earth science, engineering, health sciences, life sciences, mathematics, and physics. We calculated the proportion of invited reviewers accepting an invitation to review or re-review a manuscript per journal per month between June 1, 2001, and July 1, 2015, for single-blind and open peer review journals and between February 1, 2011, and July 1, 2015, for double-blind peer review journals.

Results The proportion of accepted invitations was 49% overall (N = 498 journals), 60% for the 40 journals implementing double-blind peer review, 53% for the 388 single-blind peer review journals, and 42% for the 70 open peer review journals (**Table 15**). Within the 4 subject areas for which data for all 3 peer review models were available—biomedicine, clinical medicine, health sciences, and life sciences journals—the proportion of reviewers accepting invitations to review was lowest for open peer review and highest for the double-blind model. A pairwise proportion test showed statistically significant differences between the proportions of reviewers who agreed to open peer review for the clinical medicine (45%), biomedicine (41%), health sciences (38%), and life sciences (31%) journals. An analysis of single-blind peer review data available for each field showed the proportion for mathematics was higher than for the other subject areas that did not exhibit differences.

Conclusions A smaller proportion of invited reviewers agreed to peer review for journals operating under open and single-blind peer review models compared with journals that use double-blind peer review. As a result, for journals with open and single-blind peer review models, a higher number of reviewers need to be invited. However, these journals have operated under their respective models for many years, so the

Table 15. Numbers of Considered Journals, Reviewer Invitations, and Manuscripts

Peer Review Model	No.					
	Manuscripts Undergoing External Review	Reviewer Invitations Sent	Reviewer Invitations Sent per Manuscript	Reviewers Who Accepted Invitations	Proportion of Accepted Invitations	Accepted Invitations per Manuscript
Double blind (n = 40)	2988	12,471	4.17	7528	0.60	2.52
Single blind (n = 388)	187,332	874,487	4.67	467,642	0.53	2.50
Open (n = 70)	101,606	598,080	5.89	254,086	0.42	2.50
All (N = 498)	291,926	1,485,038	5.09	729,256	0.49	2.50

discrepancy between these proportions does not seem to be detrimental to the success of a journal.

¹BioMed Central, London, UK, maria.kowalczyk@biomedcentral.com; ²Nature America, New York, NY, USA

Conflict of Interest Disclosures: Maria Kowalczyk is employed by BioMed Central (part of Springer Nature), and Michelle Samarasinghe is employed by Nature America (part of Springer Nature).

A Novel Open Peer Review Format for an Emergency Medicine Blog

Scott Kobner,^{1,2} Derek Sifford,^{2,3} Michelle Lin^{2,4}

Objective The medical education blog Academic Life in Emergency Medicine (ALiEM) has developed a new, open, inline peer review publication format that presents reviewer commentary within the body of the content. We hypothesized that the proximity and interactive nature of inline text will increase webpage engagement and investigated the association of this format with reader behavior.

Design We wrote 2 ALiEM blog posts for the purpose of this study. Each post was reviewed by a pair of expert peer reviewers to produce 3 publication versions: (1) control post with no expert peer review; (2) traditional post with expert peer review appended to the end of the post content; and (3) experimental post, which used inline text throughout the content. Website visitors were randomized to view either the control, traditional, or experimental version using a custom Google Analytics Content Experiment. Google Analytics, paired with a customized JavaScript activity tracker, and CrazyEgg, an industry standard web visualization suite, captured user demographic information, active users every 5 seconds, bounce rate, user click activity, and user scroll activity over a 1-month time period. We compared measures using the χ^2 test.

Results Data collection is in progress, and we report the first blog post's preliminary analytic data. During the period from June 4 to 9, 2017, 502 views were captured across the 3 versions of the post (196 views of the control post, 149 views of the traditional post, and 157 views of the experimental post). Bounce rates were nearly identical across groups (73.6%-73.9%; $P > .99$). The time users spent viewing each blog post, measured as active users present over increments of 5 seconds, was not statistically significantly different (median time of 50 seconds [interquartile range, 20-90 seconds] viewing the control post, 50 seconds [interquartile range, 20-95 seconds] viewing the traditional post, and 45 seconds [interquartile range, 20-99 seconds] viewing the experimental post; $P = .46$). There were no apparent differences in user scrolling data. However, pages featuring inline peer review comments had a statistically significantly greater proportion of clicks on summative peer review content than posts with traditional commentary (15.21% experimental vs 8.92% traditional; $P < .03$).

Conclusions The inline peer review publication system does not appear to alter reading times, bounce rates, or scrolling

activity. Readers did, however, click to open peer review content more when viewing the inline version compared with the traditional post. The preliminary data suggest that expert inline commentaries might increase the value of scientific content published online by increasing the visibility of published peer review content for readers. Future studies should examine the association of transparent, inline expert peer reviews with reader cognitive load and learning.

¹Department of Emergency Medicine, Los Angeles County and University of Southern California Medical Center, Los Angeles, CA, USA, scottkobner@gmail.com; ²Academic Life in Emergency Medicine, San Francisco, CA, USA; ³Wayne State University, Detroit, MI, USA; University of California, San Francisco at Zuckerberg San Francisco General, San Francisco, CA, USA

Conflict of Interest Disclosures: All authors are members of the not-for-profit organization Academic Life in Emergency Medicine, LLC. Dr Kobner is the New Submissions Editor, Mr Sifford is the Chief Technology Officer, and Dr Lin is the Editor in Chief and Chief Executive Officer. All are paid contractors for the organization. They are compensated for their efforts in maintaining the blog and other digital innovations. No other disclosures were reported.

Editorial and Peer-Review Process Innovations

Impact of a Change in Editorial Policy at Nature Publication Group (NPG) on Their Reporting of Biomedical Research

Malcolm Macleod,¹ for the NPQIP Collaborative Group

Objective To determine whether a change in editorial policy, including the implementation of a checklist, was associated with improved reporting of measures that might reduce the risk of bias.

Design In this before-after study, we included articles that described research in the life sciences published in Nature Publication Group (NPG) journals that were submitted after implementation of mandatory completion by authors of a checklist at the point of manuscript revision (May 1, 2013, to November 1, 2014). We compared these with articles describing research in the life sciences published in Nature journals that were submitted before May 2013. Similar articles in other journals were matched for date and topic. We investigated the change in proportion of articles published before and after May 2013 reporting 4 criteria: information on randomization, blinding, sample size calculation, and exclusions. We included 448 articles published in NPG journals (225 [50.2%] published before May 2013 and 223 [49.8%] published after) that were identified by an individual hired by the NPG for this specific task, working to a standard procedure; an independent investigator used PubMed's Related Citations feature to identify 447 similar articles with a similar topic and date of publication in other journals. We then redacted all publications for time-sensitive information and journal name. Redacted articles were assessed by 2 trained reviewers against a 74-item checklist, with discrepancies resolved by a third reviewer.

Results In total, 392 NPG articles and 353 similar articles in other publications described in vivo research. The number of NPG articles meeting all 4 criteria increased from 0 of 203 prior to May 2013 to 31 of 181 (17.1%) after (2-sample test for equality of proportions without continuity correction; $\chi^2 = 36.156$; $df = 1$; $P < .001$). There was no change in the proportion of similar articles in other publications meeting all 4 criteria (1 of 164 [0.6%] before; 1 of 189 [0.5%] after). Agreement between reviewers ranged from 72% (for “Does the manuscript describe which method of randomization was used to determine how samples/animals were allocated to experimental groups?”) to 90% (for “Does the manuscript describe how the sample size was chosen to ensure adequate power to detect a prespecified effect size?”).

Conclusions There was a substantial improvement in the reporting of measures that might reduce the risk of bias in in vivo research in NPG journals following implementation of a mandatory checklist policy, to a level that, to our knowledge, has not been previously observed in science journals. However, there remain opportunities for further improvement.

¹Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK, malcolm.macleod@ed.ac.uk

Conflict of Interest Disclosures: None reported.

Funding/Support: The study was funded by the Laura and John Arnold Foundation.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the abstract.

Assessment of Signing Peer Reviews in Principle and in Practice at Public Library of Science (PLOS) Journals

Elizabeth Seiver,¹ Helen Atkins¹

Objective To investigate the rate at which Public Library of Science (PLOS) peer reviewers at 3 medical journals chose to sign or not sign their reviews, thus revealing or not revealing themselves to the authors of an article, and whether authors' and reviewers' stated values about signing reviews matched their behavior.

Design Historical review signing data from 3 PLOS journals (*PLOS Computational Biology*, *PLOS Medicine*, and *PLOS ONE*) from mid-2013 through 2016 were analyzed. In addition, 1-click multiple choice surveys were appended to current system-generated emails, which were sent at the time of manuscript and review submission, to authors and reviewers; these surveys asked them about their general experience with, and preference for, signing reviews. The survey landing page included detailed follow-up questions about their selection, and comments were coded for qualitative analysis. The signing and survey data sets did not share unique identifiers and thus were not matched 1:1.

Results Of 451,306 total reviews analyzed, 34,561 (7.7%) were signed. This was higher for *PLOS Medicine* than for the other 2 journals (**Table 16**). Although the *PLOS Computational Biology* authors reported that they had received signed reviews (27 of 71 [38.0%]), the reviewers reported that they did not usually sign their reviews (75 of 627 [12.0%]). Combining all 3 PLOS journals, we found that 509 of 1072 authors preferred to receive signed reviews (47.5%) and that 372 of 2359 reviewers reported usually signing them (15.8%). In follow-up comments, the reviewers who reported usually signing reviews made the argument that signing their reviews improved accountability and constructiveness; the reviewers who reported usually not signing cited as their motivation the ability to be more honest and fear of retribution. Many reviewers who had not signed reviews simply stated that they had never been asked before or were not sure of the benefits. Authors who favored receiving signed reviews valued having the additional information on the reviewer's area of expertise and potentially having more open communication.

Conclusions The 3 PLOS journals represent different research communities with diverging attitudes and behavior toward signing reviews. *PLOS Medicine* maintains a strong tradition of review signing, whereas the *PLOS Computational Biology* community, despite familiarity with open review, tends to be fairly conservative about signing reviews. For journals wishing to encourage the adoption of signed reviews, it may be helpful to directly request a signature, provide signing incentives, describe the benefits of signing, or encourage reviewers to understand the author's perspective.

¹Public Library of Science, San Francisco, CA, USA, eseiver@plos.org

Conflict of Interest Disclosures: Dr Seiver and Ms Atkins are employees of PLOS.

The Role of Persistent Identifiers in the Peer Review Process: Use of ORCID

Alice Meadows¹

Objective When widely adopted and embedded in the research infrastructure, persistent identifiers (PIDs) enable interoperability between systems, reducing errors caused by manual data entry and saving researchers time. This study presents initial results from 2 analyses of PID adoption and use of ORCID (Open Researcher and Contributor Identifier). The first is an analysis by the Technical and Human Infrastructure for Open Research (THOR), a European Union-funded ORCID partner organization, of the uptake of ORCID identifiers. The second analysis reviews uptake of the peer review functionality associated with ORCID identifiers, which was introduced in October 2015.

Design The THOR study analyzes uptake of PIDs across 4 broad disciplines (20 subdisciplines) and 5 geographic regions to develop a baseline comparator using publicly available data on the location and journal publications of ORCID registrants and Science Metrix's subject fields. The

Table 16. Data on Reviews Signed at Each Journal

Journal	Reviews Signed, No./Total No. (%)				
	2013	2014	2015	2016	Total
<i>PLOS ONE</i>	3272/48,358 (6.8)	10,607/139,346 (7.6)	9849/134,800 (7.3)	8327/114,771 (7.3)	32,055/437,275 (7.3)
<i>PLOS Medicine</i>	159/437 (36.4)	495/1348 (36.7)	508/1280 (39.7)	533/1603 (33.3)	1695/4668 (36.3)
<i>PLOS Computational Biology</i>	NA	73/1055 (6.9)	300/4234 (7.1)	338/4074 (8.3)	711/9363 (7.6)
Total	3431/48,795 (7.0)	11,175/141,749 (7.9)	10,657/140,314 (7.6)	9198/120,448 (7.6)	34,461/451,306 (7.6)
	Survey Questions, No./Total No. (%)				
	How often do you sign your reviews? [reviewers]	Do you prefer signed reviews? [authors]	Have you received signed reviews? [authors]	Have you received signed reviews? [reviewers]	
<i>PLOS ONE</i>	257/1467 (17.5) Usually sign	486/1038 (46.8) Prefer signed	630/2108 (29.9) Yes	407/1751 (23.2) Yes	
<i>PLOS Medicine</i>	40/265 (15.1) Usually sign	8/11 (80.0) Prefer signed	17/34 (50.0) Yes	28/65 (43.1) Yes	
<i>PLOS Computational Biology</i>	75/627 (12.0) Usually sign	15/23 (65.2) Prefer signed	27/71 (38.0) Yes	50/123 (40.7) Yes	
Total	372/2359 (15.8) Usually sign	509/1072 (47.5) Prefer signed	674/2213 (30.5) Yes	485/1939 (25.0) Yes	
	Emails sent, No.				
<i>PLOS ONE</i>	7209	4907	8369	11,620	
<i>PLOS Medicine</i>	1569	354	879	416	
<i>PLOS Computational Biology</i>	3985	182	394	965	
	Response Rate, %				
<i>PLOS ONE</i>	20	21	25	15	
<i>PLOS Medicine</i>	17	3	4	16	
<i>PLOS Computational Biology</i>	16	13	18	13	

Abbreviations: PLOS, Public Library of Science; NA, not available.

preliminary results of this analysis provided context for the study on uptake of ORCID's peer review functionality. From October 2015 to the end of May 2017, 135,605 review activities were added to 9803 ORCID records by 9 organizations. Focusing on data from 3 early adopters (Publons, the American Geophysical Union, and *F1000*, collectively representing 99.59% of total use), use to date was assessed, including the number of reviewers whose ORCID records contain review activities, whether this information is publicly available, the number of review activities added (including number of DOIs where applicable), and number of participating journals. This information was supplemented by informal feedback from reviewers about the functionality and its benefits and drawbacks.

Results The discipline-only analysis of ORCID (487,471 ORCID records; 3,703,958 publications) showed the top share of ORCID in clinical medicine (507,230 publications [13.7%]), technology and other applied sciences (468,676 publications [12.7%]), and biology (454,468 publications [12.3%]). Built and environmental design had the least uptake (9800 publications [0.3%]). The regional-only analysis (785,020 ORCID records) showed the most uptake of ORCID in Europe (326,136 records [41.5%]) and the least uptake in the Middle East and Africa (36,768 records [4.7%]). In terms of ORCID's peer review functionality, Publons is the top user. Of the 151,973 Publons users as of the end May 2017, 10,471 (6.89%) have connected a total of 125,892 review activities to

their ORCID records. This represents 92.8% of all review activities in ORCID. In addition, *F1000* has connected 5714 reviews (4.21%) to ORCID records and the American Geophysical Union has connected 3455 reviews (2.55%). The number of participating journals from these 3 organizations was 8763. Initial feedback from participating reviewers indicates that they find the functionality of linking reviewer activity to ORCID records valuable. However, there is a low level of knowledge and understanding of the option to connect review activities to ORCID records among reviewers, journals, and their organizations, indicating a need for increased outreach and education.

Conclusions Use of PIDs—both in general and for peer review activity—varies by discipline and country. If widely adopted in a variety of peer review workflows in the future, ORCID could help address issues around recognition for peer review in all its forms.

*Community Engagement and Support, ORCID, Brookline, MA, USA, a.meadows@orcid.org

Acknowledgments: Ms Meadows is employed by ORCID and acknowledges the assistance of ORCID colleagues and from the THOR Project, as well as Brooks Hanson and Jeanette Panning of the American Geophysical Union, Rebecca Lawrence of *F1000*, and Tom Culley and Andrew Preston of Publons.

Conflict of Interest Disclosures: None reported.

Introduction of Patient Review Alongside Traditional Peer Review at a General Medical Journal (*The BMJ*): A Mixed Methods Study

Sara Schroter,¹ Amy Price,^{1,2} Rosamund Snow,^{1,3} Tessa Richards,¹ Sam Parker,¹ Elizabeth Loder,¹ Fiona Godlee¹

Objective To evaluate the feasibility of incorporating patient reviews into the traditional peer review process at *The BMJ*.

Design This was a mixed methods study including a comparison of acceptance and completion rates and timeliness to review between patient reviewers and traditional reviewers for articles sent for peer review in 2016. We also surveyed the patient reviewers and research editors on their views of the value of patient reviews.

Results In 2016, 359 of 647 research articles (55%) sent for review had at least 1 patient reviewer invitation. For review invitations in 2016, the agreement rate for patient reviewers was 287 of 677 (42%) and the completion rate was 224 of 287 (78%); for traditional reviewers, the agreement rate was 2649 of 6998 (38%) and the completion rate was 2217 of 2649 (84%). Patient reviewers took a mean (SD) of 10.5 (5.9) days to complete a review after agreement compared with 13.4 (7.7) days for traditional reviewers. Overall, 122 of 164 patient reviewers (74%) responded to a survey, and 100 of those patient reviewers (82%) would recommend being a patient reviewer for *The BMJ* to other patients and carers. One hundred seven of the patient reviewers who responded to a survey (88%) think more journals should adopt patient review, and 98 (80%) did not have any concerns about doing open review. Of the 20 patient reviewers who reviewed papers that were returned to the authors for revisions, 15 agreed or strongly agreed that the authors addressed their points, and 15 agreed that the authors were courteous when addressing their points. Twelve patient reviewers who reviewed papers that were returned to the authors for revisions felt they included points important to patients that were not raised by the traditional reviewers. Seven of 8 research editors responded to the editor survey; 5 of 7 reported patient reviews currently add “a little” value to research papers; and 2 of 7 believed patient reviews add “a lot” of value to research papers. However, 5 of 7 research editors found it difficult to identify appropriate patient reviewers, and 5 of 7 experienced difficulty communicating with patient reviewers about articles. All editors reported patient reviewers “occasionally” include insights not raised by other reviewers; 6 of 7 editors “occasionally” and 1 of 7 editors “frequently” find patient reviewers’ comments helpful when advising authors on revisions to manuscripts. Four editors felt that other journals should adopt patient review, and 3 were unsure.

Conclusions Patient review of research is feasible alongside a standard peer review process and is considered beneficial by some editors and important by patients and carers. Further qualitative research should capture the value of the changes made to manuscripts as a result of patient reviews.

¹*The BMJ*, London, UK, sschroter@bmj.com; ²Department of Continuing Education, University of Oxford, Oxford, UK; ³University of Oxford Medical School, Oxford, UK

Conflict of Interest Disclosures: Dr Schroter, Dr Richards, Mr Parker, Prof Loder, and Dr Godlee are employees of *The BMJ*; Dr Snow was also an employee of *The BMJ*. Dr Godlee is a member of the Peer Review Congress Advisory Board but was not involved in the review or decision for this abstract. Dr Price is a Research Fellow at *The BMJ*.

Acknowledgments: We thank all the patient reviewers for completing the survey and all those who have given their time to review articles for *The BMJ* and helped us improve our guidance. We also thank *The BMJ* editors for embedding patient review into their daily work flow.

Additional Information: Dr Rosamund Snow is deceased.

Prepublication and Postpublication Issues

Associations Between bioRxiv Preprint Noncitation Attention and Publication in the Biomedical Literature: A Cross-sectional and Cohort Study

Stylianos Serghiou¹, John P. A. Ioannidis¹⁻⁴

Objective To describe associations between bioRxiv preprint traffic, Altmetric scores, and eventual publication and to compare the attention ex-preprints receive when published in the canonical literature with the attention given to published articles not republished on bioRxiv.

Design We downloaded all preprints available on bioRxiv through January 17, 2017; all metrics available for each preprint; all data held for each preprint by Altmetric; and all data held by Altmetric and PubMed for published articles of previous preprints, which are identified in bioRxiv. We randomly chose 211 published articles that had been bioRxiv preprints, randomly identified 5 journal and time-matched control articles that had not been preprinted on bioRxiv, and compared Altmetric data. We compared means using pairwise *t* and Wilcoxon signed rank tests, estimated associations between the preprint covariable for the field of study and canonical publication using a multivariable Cox proportional hazards model, and compared matched data using a mixed-effects model with random intercept.

Results Of 7760 preprints, median traffic to abstracts was 943 (range, 6-192,570) and median traffic to PDFs was 331 (range, 16-151,520). Median Altmetric score was 7.3 (range, 0.25-2506) with a heavy right skew. Two thousand thirty-one preprints (36%) reached the canonical literature within a year of being uploaded and had a higher mean Altmetric score (19.5 vs 11.4, $P = 3.8 \times 10^{-8}$) than articles that had not reached the canonical literature within a year. After adjusting for the field of study, the Altmetric score remained a statistically significant but weak variable associated with eventual publication (hazard ratio, 1.005; 95% CI, 1.002-1.007). Once a preprint article is published, the pairwise absolute mean difference in Altmetric score was 14.9 points higher than what it was as a preprint on bioRxiv ($P = 10^{-5}$). The biggest

contributor to this difference was the number of citations by the mainstream media (pairwise absolute mean [SE] difference, 16.6 [3.8]). Articles that had previously been posted on bioRxiv attracted significantly more attention than controls once they were published (mean difference, 19.2; 95% CI, 5.5-32.8).

Conclusions Many preprints on bioRxiv attract significant noncitation attention and reach canonical publication. The Altmetric score is not meaningfully associated with eventual publication, but articles that had been posted on bioRxiv tend to receive more attention once published than articles that had never been posted on bioRxiv.

¹Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA, USA; ssteli@stanford.edu;
²Stanford Prevention Research Center, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA;
³Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, CA, USA; ⁴Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA, USA

Conflict of Interest Disclosures: Dr Ioannidis is a Peer Review Congress Associate Director but was not involved in the review or decision for this abstract.

Differences in Readership Metrics and Media Coverage Among Negative, Positive, and Mixed Studies Published by the *New England Journal of Medicine*

Ramya Ramaswami,¹ Sagar Deshpande,^{2,3} Rebecca Berger,¹ Pamela Miller,¹ Edward W. Campion¹

Objective Negative studies are defined as reports where there is no statistical difference between groups in the primary outcome. These studies may not be published owing to several factors, including hesitation by authors to submit negative studies for publication. When negative studies do proceed to publication, it is unclear how much attention they receive from readers and the media. We analyzed whether there were differences in readership metrics and media coverage among negative, positive, and mixed studies published by the *New England Journal of Medicine* (NEJM).

Design NEJM.org tracks and displays metrics on readership and media coverage using 3 online analytic sources: Atypion.com, Crossref, and Cision. We retrieved information on page views (number of times the article was accessed online), citations (number of citations by peer-reviewed journals), and media coverage (number of unique media mentions) for all reports of clinical trials published in NEJM between 2012 and 2015. Readership and media coverage metrics were collected from date of publication through January 2017. The papers were labeled as negative, positive, or mixed (discordant coprimary end points) based on results for the study end points. Readership metrics and media coverage means for the 3 groups were assessed by analysis of variance (ANOVA), and adjustments for multiple comparisons were made using the Scheffé method.

Results A total of 338 articles were included in the analysis, of which 73 (22%) were negative studies, 224 (66%) were positive, and 41 (12%) were mixed. There was 100% agreement on classification of articles by 2 authors (S.D. and P.M.). For the 73 negative studies, the article metrics were as follows: page views, 58,728; citations, 93; media coverage, 101. For the 224 positive studies, the article metrics were as follows: page views, 64,364; citations, 88; media coverage, 120. For the 41 mixed studies, the article metrics were as follows: page views, 43,810; citations, 62; media coverage, 65. There were no statistically significant differences across the 3 groups in mean page views, mean citations, and mean media coverage (Table 17). Following adjustment for multiple comparisons, there were no statistically significant differences in readership or media coverage metrics among the 3 groups. A larger study would be required to assess the generalizability of these findings to other journals and to evaluate other factors that influence postpublication metrics.

Conclusions There was no difference in mean page views, citations, and media coverage among positive trials, negative trials, and trials with mixed outcomes published at NEJM.org.

¹New England Journal of Medicine, Massachusetts Medical Society, Boston, MA, USA, rramaswami@nejm.org; ²University of Michigan Medical School, Ann Arbor, MI, USA; ³John F. Kennedy School of Government at Harvard University, Cambridge, MA, USA

Conflict of Interest Disclosures: Dr Campion is a member of the Peer Review Congress Advisory Board but was not involved in the review or decision for this abstract.

Funding/Support: Drs Ramaswami, Berger, and Campion, and Ms Miller receive salary support from the Massachusetts Medical Society.

Table 17. Mean Page Views, Citations, and Media Coverage in Negative, Positive, and Mixed Studies

Variable	73 Negative (95% CI)	224 Positive (95% CI)	41 Mixed (95% CI)	Unadjusted P Value
Page views	58,728 (57,707- 59,749)	64,364 (63,712- 65,015)	43,810 (42,408- 45,212)	.16
Citations	93 (89-96)	88 (87-90)	62 (58-65)	.47
Media coverage	101 (97-105)	120 (118-121)	65 (61-69)	.09

Reproducible Research Practices in Systematic Reviews of Therapeutic Interventions: A Cross-sectional Study

Matthew J. Page,¹ Douglas G. Altman,² Larissa Shamseer,^{3,4} Joanne E. McKenzie,¹ Nadera Ahmadzai,⁵ Dianna Wolfe,⁵ Fatemeh Yazdi,⁵ Ferrán Catalá-López,^{5,6} Andrea C. Tricco,^{7,8} David Moher^{3,4}

Objective Biomedical researchers are increasingly encouraged to use reproducible research practices, which allow others to recreate the findings, given the original data. Such practices include providing a detailed description of the data collected and used for analysis, clearly reporting the analysis methods and results, and sharing the data set and

statistical code used to perform analyses (within the journal article, a supplementary file, or a data repository). To our knowledge, there has been no investigation into how often such practices are used in systematic reviews (SRs) across different specialties. We aimed to investigate reproducible research practices used in a cross-section of SRs of therapeutic interventions.

Design We selected articles from a database of SRs we assembled previously, which included a random sample of 300 SRs that were indexed in MEDLINE during February 2014. In the current study, we included only those SRs that focused on a treatment or prevention question and reported at least 1 meta-analysis. One author collected data on 28 prespecified items that characterized reproducible research practices from the SR article and any supplementary files; a 20% random sample was collected in duplicate. We did not contact authors of the SRs for additional information. We calculated risk ratios to explore whether reproducible research practices differed between Cochrane and non-Cochrane SRs.

Results We evaluated 110 SRs; 78 (70.9%) were non-Cochrane articles, and 55 (50.0%) investigated a pharmacological intervention. The SRs presented a median (interquartile range) of 13 (5-27) meta-analyses. Authors of SRs reported the data needed to recreate all meta-analytic effect estimates in the SR, including subgroup meta-analytic effects and sensitivity analyses, in only 72 of 110 SRs (65.5%). This percentage was higher in Cochrane than in non-Cochrane SRs (94% vs 54%) (**Figure 7**). Despite being recommended by PRISMA, summary statistics (eg, means and SDs) of each individual study were not reported for 31 of 110 index (ie, primary or first-reported) meta-analyses (28.2%). Authors of SRs who reported imputing, algebraically manipulating, or obtaining some data from the included studies' authors/sponsors infrequently stated which specific data were handled in this way. Only 33 SRs (30.0%) mentioned access to data sets and statistical code used to perform analyses.

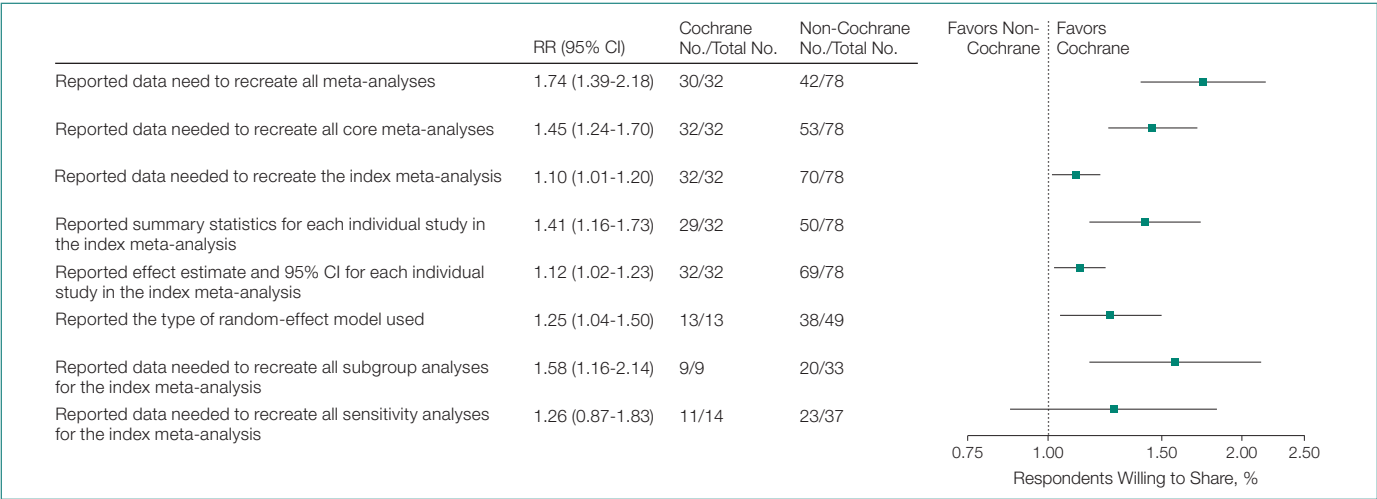
Conclusions Reproducible research practices in SRs of therapeutic interventions are suboptimal. Authors of SRs should make greater use of public data repositories (eg, the Systematic Review Data Repository or Open Science Framework) to share SR data sets and statistical analysis code so that others can recreate the findings, check for errors, or perform secondary analyses.

¹School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia, matthew.page@monash.edu; ²UK EQUATOR Centre, Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences, University of Oxford, Oxford, UK; ³Centre for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada; ⁴School of Epidemiology, Public Health, and Preventive Medicine, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada; ⁵Knowledge Synthesis Group, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada; ⁶Department of Medicine, University of Valencia/INCLIVA Health Research Institute and CIBERSAM, Valencia, Spain; ⁷Knowledge Translation Program, Li Ka Shing Knowledge Institute, St Michael's Hospital, Toronto, Ontario, Canada; ⁸Epidemiology Division, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada

Conflict of Interest Disclosures: Matthew Page and Joanne McKenzie are affiliates of Cochrane Australia. Matthew Page is a coconvenor of the Cochrane Bias Methods Group. Joanne McKenzie is a coconvenor of the Cochrane Statistical Methods Group. Andrea Tricco is an author of 2 of the systematic reviews included in this study but was not involved in eligibility assessment or data collection. Drs Altman and Moher are members of the Advisory Board for the Peer Review Congress but were not involved in the review or decision for this abstract.

Funding/Support: Matthew Page is supported by grant 1088535 from the Australian National Health and Medical Research Council Early Career Fellowship. Douglas Altman is a National Institute for Health Research Senior Investigator. Joanne McKenzie is supported by grant 1072366 from the Australian National Health and Medical Research Council Australian Public Health Fellowship. Ferrán Catalá-López is supported by the Generalitat Valenciana (PROMETEOII/2015/021). Andrea Tricco is funded by a Tier 2 Canada Research Chair in Knowledge Synthesis. David Moher is supported in part by a University Research Chair, University of Ottawa.

Figure 7. Risk Ratio (RR) Associations Between Type of Systematic Review (Cochrane vs Non-Cochrane) and Reproducible Research Practices



Postpublication Issues

Analysis of Indexing Practices of Corrected and Republished Articles in MEDLINE, Web of Science, and Scopus

Tea Marasović,¹ Ana Utrobičić,^{2,3,4} Ana Marušić^{3,4}

Objective Recently updated International Committee of Medical Journal Editors (ICMJE) recommendations suggest correcting honest errors by “retraction with republication of the changed paper, with an explanation.” MEDLINE uses “Corrected and Republished Article” to indicate correction of “a previously published article by republishing the article in its entirety.” We assessed how other bibliographical databases indexed article corrections with replacement.

Design Articles indexed as “Corrected and Republished Articles” in MEDLINE from January 2015 to December 2016 (n = 29) were analyzed for the information presented in journals and in the Web of Science (WoS) and Scopus. Two authors independently extracted the data and reached a consensus for disagreements.

Results Twenty-nine articles were published in 24 biomedical journals from different research areas (all articles were indexed in Scopus and all but 3 indexed in WoS; median Impact Factor for WoS journals, 2.98; 95% CI, 2.15-3.35). Half of the journals published a separate item to indicate correction, and half had a CrossMark tag on the corrected article (**Table 18**); CrossMark tags had no links to the original article or notification of correction. PubMed did not provide the links between the corrected to the original article in 1 case. Web of Science and Scopus indexed corrected articles most often as a correction (WoS) or erratum (Scopus). Five articles in WoS (17%) and 11 in Scopus (38%) were indexed as “articles” in the same way as original articles, which made it difficult to differentiate between the versions. When corrected articles were indexed, they often lacked links to the original articles. Original and corrected articles had a similar median number of citations (WoS: original, 2.0; 95% CI, 2.0-4.2; corrected, 2.0; 95% CI, 1.0-6.8; Scopus: original, 2.0; 95% CI, 1.0-4.4; corrected, 2.0; 95% CI, 1.0-10.2). A PubMed search for “retraction and replacement” identified 5 more articles published in *JAMA*, which were indexed only as retracted and not as corrected and republished publications. Web of Science indexed only 1 version of these articles, and the notices of retraction and replacement were indexed as a “letter,” “correction,” or “editorial material.” Scopus also indexed all 5 articles, and notices were indexed as a “letter” or “erratum” (1 was missing).

Conclusions There seem to be serious discrepancies in indexing corrected and republished articles in major databases, which diminishes the credibility and transparency of the research and publication system. While scientific self-correction should be supported, all stakeholders in the publication process should commit to ensuring that published scientific articles are appropriately indexed, interlinked, updated, and/or amended in a timely and efficient manner.

Table 18. Characteristics of 29 Articles Indexed by MEDLINE as “Corrected and Republished Article” from 2015 to 2016 and their Indexation in Web of Science and Scopus

Characteristic	Journal ^a	WoS	Scopus
Corrected article published as separate article	25	NA	NA
Separate item published to indicate correction	15	NA	NA
Correction indicated in the republished article	15	NA	NA
CrossMark tag present	15	NA	NA
Original article indexed	NA	23	22 ^b
Corrected article indexed as	NA	22	25 ^b
Article	NA	5	11
Article; retracted publication	NA	2	NA
Correction (WoS/Scopus)	NA	11	1
Erratum (Scopus)	NA	NA	11
Editorial material	NA	2	NA
Reprint	NA	1	NA
Review	NA	1	1
Note	NA	NA	1
Corrected article links to original article	NA	12	8 ^c
Separate item indicating correction indexed as	NA	9	10
Correction (WoS) or erratum (SCOPUS)	NA	9	9
Note	NA	NA	1
Separate item indicating correction links to original article	NA	9	10

Abbreviations: NA, not applicable; WoS, World of Science.

^aFull text of 1 article could not be obtained, so 28 articles were used for analysis in this category.

^bOne article was indexed twice (considered as a single indexation in this table).

^cIn 1 case, only the link to the journal title was provided, without other identifying article elements.

¹Research Office, University of Split School of Medicine, Split, Croatia; ²Central Medical Library, University of Split School of Medicine, Split, Croatia; ³Cochrane Croatia, Split, Croatia, ana.marusic@mefst.hr; ⁴Department of Research in Biomedicine and Health, University of Split School of Medicine, Split, Croatia

Conflict of Interest Disclosures: Dr Marušić is a member of the Peer Review Congress Advisory Board but was not involved in the review or decision for this abstract.

Funding/Support: This study was funded by grant IP-2014-09-7672 from the Croatian Science Foundation.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the abstract.

A Cross-sectional Study of Commenters and Commenting in PubMed, 2014 to 2016: Who's Who in PubMed Commons

Melissa D. Vaught,¹ Diana C. Jordan,¹ Hilda Bastian¹

Objective Authors indexed in PubMed are eligible to join PubMed Commons and post English-language comments that appear directly below abstracts in PubMed. Journal club membership was introduced in 2014. We sought to describe

characteristics of commenters and the extent of collaborative and author comments and replies.

Design Basic usage data were collected for comments posted from January 2014 to December 2016 (online February 2017). We evaluated a 6-month subset for collaborative and author posts, as well as for commenters' gender, country, and conflict of interest disclosure.

Results At the close of 2016, 10,736 individuals and 24 journal clubs had joined PubMed Commons. From 2014 to 2016, 5483 comments were posted to 4372 publications, with 13% of individuals (n=1410) and 71% of journal clubs (n=17) commenting. The mean (range) number of comments per active individual was 4 (1-196), with 38% (n=537) posting more than 1 comment. For active journal clubs, the mean (range) was 8 (1-27), with 82% (n=14) posting multiple comments. From July 2016 to December 2016, 953 comments (17% of the 3-year total) were posted to 776 publications (18% of the 3-year total) by 332 members and a further 69 named coauthors (**Table 19**). Commenters were primarily from 5 English-speaking countries (n=244 [63%]) and 21% of all commenters were women. Authors posted 71 replies (8%) and 109 other comments (11%). Collaborative comments accounted for 21% of posts, including multiauthored and collective (eg, journal club) comments. Conflict of interest disclosures were formally or informally disclosed rarely (23 instances), often to declare an absence of conflicts of interest.

Conclusions Most individual members of PubMed Commons have not commented, although a small number of members account for a considerable proportion of comments. Comments rarely include conflict of interest disclosures. Geographic distribution of commenters is not representative of authors in the biomedical literature, and women are underrepresented. Author replies are uncommon. Many comments are collaborative posts.

Table 19. PubMed Commons Commenter Characteristics, July to December 2016

Characteristic	No. (%)
PubMed Commons Journal Clubs	8
Individual PubMed Commons members	324
Additional named coauthors ^a	69
Geographic location of individual members and named coauthors ^b	389
Europe	168 (43)
North America	163 (42)
Asia or Oceania	38 (10)
Other	20 (5)
Unknown	4 (1)
Gender of individual members and named coauthors ^b	387
Male	304 (79)
Female	83 (21)
Unknown	6 (1)

^aIndividuals listed as coauthor of a comment posted from an individual PubMed Commons account.

^bFor individual PubMed Commons members and named coauthors (n = 393), geographic location and gender were determined.

PubMed Commons, National Center for Biotechnology Information (NCBI), US National Library of Medicine (NLM), National Institutes of Health (NIH), Bethesda, MD, USA, melissa.vaught@nih.gov

Conflict of Interest Disclosures: All authors worked at NCBI when this study was being conducted, including when data were collected from PubMed Commons.

Funding/Support: This work was supported by the Intramural Research Program of the NIH, NLM.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the abstract.

The Role of PubPeer Comments in Alerting Editors to Serious Problems With Clinical Research Publications

Elizabeth Wager,^{1,2} Emma Veitch³

Objective PubPeer is a self-described “online journal club” that facilitates commenting on published biomedical literature. We sought to determine how often postpublication comments on PubPeer identify serious misconduct or errors in clinical research articles, how often editors are alerted to problems via PubPeer, and how editors and authors respond.

Design Two raters independently categorized all comments on PubPeer about research publications in *BMC Medicine*, *The BMJ*, and *The Lancet* from first comment appearance (October 2013) to December 31, 2016 (comments on editorials, letters, news, etc were counted but not analyzed). The categories, developed iteratively and by consensus, included well-supported allegations of fabrication, falsification, or plagiarism (FFP); vague FFP allegations (presenting no evidence); allegations of other misconduct; honest error; and methodological concerns. Differences were resolved by discussion. We contacted editors to ask whether PubPeer alerted them to the allegations and how they responded.

Results We found 344 PubPeer comments relating to 150 articles. Of 177 comments relating to 99 research articles, 106 (60%) were imported from PubMed Commons (PMC) (all signed, as required by PMC), of which 11 (6%) were from journal clubs. Of the non-PMC comments, 67 (94%) were anonymous. Of the 177 comments on research articles, 7 (4%; 2 signed) made allegations about or mentioned investigations into FFP in 4 articles (3 strong, 4 vague), 5 (3%; 4 signed) identified errors in 5 articles (mainly concerning trial registration identifier numbers), 29 (16%; 26 signed) raised methodological issues about 20 articles, and 16 (9%) discussed clinical implications. Fifty-nine comments (33%) contained little or no text but gave links to other sites (eg, journal articles, blogs, retraction notices), and 10 (6%) provided extra information without criticism. Journal editors were unaware of the PubPeer postings about their published articles but had independently issued corrections (3) or expressions of concern (2). Authors responded on PubPeer to comments about 4 articles (4%). Commentary on other types of research (eg, comments on basic science, which occur more

frequently on PubPeer than comments on clinical studies), on other sites, and other editors' responses may be different from that on PubPeer.

Conclusions Only 7% of comments on 9 research articles in our sample raised issues that might require journal action (7 fraud, 5 error). The 3 journals had not been alerted to problems via PubPeer but were generally aware of the concerns from other sources and issued corrections (3), or expressions of concern (2). While PubPeer provides a useful forum for postpublication comments, the frequency of comments requiring journal action in our clinical journal sample was low.

¹Sideview, Princes Risborough, UK, liz@sideview.demon.co.uk;

²University of Split Medical School, Split, Croatia; ³Freelance Editor, London, UK

Conflict of Interest Disclosures: None reported.

Funding/Support: This study was funded by Sideview, which is owned by Elizabeth Wager and paid Emma Veitch for her work on the study.

Acknowledgments: We thank Jigisha Patel and Lin Lee (*BMC Medicine*), Theodora Bloom (*The BMJ*), and Sabine Kleinert (*The Lancet*) for supplying information about journal editors' awareness of PubPeer comments. We thank Brandon Stell of PubPeer for answering queries about the mechanisms of PubPeer.

Poster Session Abstracts

Posters will be presented during 1 of 2 sessions, on Monday, September 11, and Tuesday, September 12.

Authorship and Contributorship

MONDAY

Trends in Authorship and Team Science in Major Medical Journals, 2005-2015

Christopher C. Muth,¹ Robert M. Golub¹

Objective Team science helps address complex research questions by encouraging interdisciplinary and multicenter collaborations. Recognizing the value of team science but also the importance of acknowledging individuals' contributions, journals may include increased numbers of authors on a byline, group authorship, and special designations to indicate authors who contributed equally to the work when publishing research. This study assessed trends in authorship of research articles published in 3 major medical journals to test the hypothesis that team science is supported by major medical journals and has increased over time.

Design Research articles published in 2005, 2010, and 2015 in the top 3 general medical journals (*JAMA*, *Lancet*, and *New England Journal of Medicine* [*NEJM*]) based on current Impact Factor were identified using the Web of Science database. Reviews and meta-analyses were excluded. The number of authors, presence of group authorship, and presence of authors who contributed equally (detected by manual review of the byline for asterisks or other notation and manual assessment of the endnotes for relevant statements) were determined for each article. Trends in these authorship metrics were then assessed by journal over time.

Results The numbers of articles that met inclusion criteria were 230, 188, and 159 in *JAMA*; 172, 165, and 178 in *Lancet*; and 223, 222, and 235 in *NEJM* for the years 2005, 2010, and 2015, respectively. The median number of authors per article and the proportion of articles with authors who contributed equally increased significantly over time for all journals (**Table 20**). The proportion of articles with group authorship increased significantly over time for *JAMA* but not for *Lancet* or *NEJM*.

Conclusions The number of authors per article and the proportion of articles with authors who contributed equally increased over time in 3 major medical journals. Although limited to top general medical journals, these findings are consistent with previous studies focused on earlier periods and specialty journals. Increases in these authorship metrics may indicate an increase in team science and suggest that major medical journals reflect this trend.

Table 20. Authorship Trends by Journal Over Time

	Year			P Value for Trend
	2005	2010	2015	
Authors per Article, Median (IQR)				
JAMA	8 (5-11)	8 (6-12)	11 (7-18)	<.001
Lancet	9 (7-13)	12 (8-16)	15 (10-21)	<.001
NEJM	11 (7-15)	13 (9-20)	18 (12-26)	<.001
Articles With Group Authorship, No./Total (%)				
JAMA	38/230 (16.5)	40/188 (21.3)	47/159 (29.6)	.002
Lancet	66/172 (38.4)	67/165 (40.6)	75/178 (42.1)	.47
NEJM	82/223 (36.8)	87/222 (39.2)	106/235 (45.1)	.07
Articles With Authors Who Contributed Equally, No./Total (%)				
JAMA	7/230 (3.0)	13/188 (6.9)	17/159 (10.7)	.002
Lancet	9/172 (5.2)	16/165 (9.7)	31/178 (17.4)	<.001
NEJM	22/223 (9.9)	25/222 (11.3)	64/235 (27.2)	<.001

Abbreviations: IQR, interquartile range; *NEJM*, *New England Journal of Medicine*.

¹*JAMA*, Chicago, IL, USA, christopher.muth@jamanetwork.org

Conflict of Interest Disclosures: The authors are employed at *JAMA*, one of the journals included in this study. There are no other conflicts of interest to report.

Funding/Support: There was no external funding for this study.

Additional Contributions: Joseph Wislar, a former employee at *JAMA* and currently employed at DentaQuest Institute, Williamsport, Pennsylvania, provided statistical consultation.

TUESDAY

Frequency of Reporting on Patient Involvement in Research Studies Published in a General Medical Journal: A Descriptive Study

Amy Price,^{1,2} Sara Schroter,¹ Rosamund Snow,^{1,3} Sophie Staniszewska,^{4,5} Sam Parker,¹ Tessa Richards¹

Objective The requirements for planning of public involvement in research—ie, research “with” or “by” members of the public rather than “to,” “about,” or “for” them—within grant applications has increased. To date, there is not an agreed method of reporting public involvement in research, and this can make such involvement challenging to identify. To address this, *The BMJ* now asks submitting authors to include a dedicated section on how they involved patients in their research and, if they did not, to state there was no involvement. We explore the early influence on public involvement reporting, frequency, and practice following the introduction of a mandatory public involvement section.

Design We report a before-and-after comparison of published research articles to assess whether the rate of reporting of public involvement in research increased with the introduction of a mandatory section for describing this involvement. Two researchers independently extracted data and reached consensus on incidences and types of public involvement in research across two 12-month samples. No study designs were excluded, because public involvement in research is possible with studies that have no direct contact with participants, eg, systematic reviews.

Results Between June 1, 2013, and May 31, 2014, *The BMJ* published 189 research articles. Two (1.1%) reported public involvement activity. From June 1, 2015, to May 31, 2016, following the introduction of the public involvement section, *The BMJ* published 152 research articles, of which 16 (10.5%) reported public involvement. Patients were included in multiple aspects of research, from grant applications and study design to coauthorship and dissemination (**Table 21**). Of the 18 articles including some information on public involvement, 6 (33.3%) clearly acknowledged patients' help or commented on the value of their contributions, and 2 (11.1%) included patient contributors as coauthors.

Conclusions Public involvement in research is not commonplace, despite being encouraged by research funders. This is not solely a reporting issue, as the proportion of papers reporting public involvement was modest, even after introducing the mandatory public involvement declaration within the methods section of *The BMJ* articles. Some authors may have initiated their research prior to *The BMJ* mandatory public involvement reporting initiative, but some ethical review boards and funding agencies have been requesting this involvement for some years. Journals and funders should collaborate to improve guidance on how to involve and report patient involvement in research. Reporting innovative ways patients are involved in research processes may encourage practice in this important area.

¹*The BMJ*, London, UK, dr.amyprice@gmail.com; ²Department of Continuing Education, University of Oxford, Oxford, UK; ³University of Oxford University Medical School, Oxford, UK; ⁴Patient and Public Involvement and Experiences of Care Program, University of Warwick, Coventry, UK; ⁵*Research Involvement and Engagement*

Conflict of Interest Disclosures: Sara Schroter, Rosamund Snow, Sam Parker, and Tessa Richards are employed by *The BMJ*, which has a patient partnership initiative. Amy Price, Rosamund Snow, and Tessa Richards are patients with long-term medical conditions and are committed to the involvement of patients in all stages of the research process. Amy Price is a research fellow at *The BMJ*. Sophie Staniszewska has no conflicts of interest.

Funding/Support: We received no external funding for this study.

Additional Contributions: Our powerful and inspiring coauthor, Rosamund Snow, passed away before this work could be published. We are thankful for her flexibility and steadfast leadership as we launched this initiative. We thank Sarah Foster (intern, *The BMJ*) for her help with the project.

Table 21. Type of Patient Involvement Explicitly Reported Before and After the Introduction of a Mandatory Section^a

Involvement Type	No. (%)	
	Preimplantation (n=2)	Postimplantation (n=16)
Study Design		
Contributions to the grant application and/or study protocol	0	3 (19)
Help setting the research question or commenting on its importance	0	4 (25)
Ensuring the development of, or choice of, outcome measures was informed by patients' priorities, experience, and preferences	1 (50)	7 (44)
Study Conduct		
Involved in the study steering group or a member of the research team	1 (50)	6 (38)
Recruitment and/or implementation of the research	1 (50)	8 (50)
Patient/public communication materials, eg, patient information sheets	0	4 (25)
Analysis		
Contributed to data analysis	0	1 (6)
Interpretation of study findings	0	2 (13)
Drafting of the Manuscript		
Contributions to the editing, revising, or writing of the manuscript	1 (50)	3 (19)
Patients listed as coauthors	1 (50)	1 (6)
Dissemination		
Direct involvement of patients led by the research team, including the development of materials for dissemination and choosing the most appropriate method of dissemination	0	3 (19)
Indirect involvement through dissemination to patient charities, organizations, and groups that may in turn involve patients in the process	0	3 (19)
Patient representation informing the content of dissemination materials on a general advisory board for the use of the data used in research	0	3 (19) ^b

^aResponses are not mutually exclusive. For example, if a patient was on the steering group, this will be indicated in the relevant box and in the box about implementation of the research. However, not all those involved in study conduct were made members of steering groups.
^bIncludes 3 articles from the QResearch team with identical statements about some aspects of public involvement in dissemination, but this public involvement was specific to the QResearch database, not the individual published studies, and it is not clear how much patients were involved in the individual studies reported.

MONDAY

Authorship for Sale: A Survey of Predatory Publishers and Journals

Pravin M. Bolshete¹

Objective To study the attitude of possible or probable predatory publishers and journals towards unethical requests to add authors.

Design Survey study of predatory publishers and stand-alone journals (a publisher that publishes only 1 journal) between November 2015 and December 2016. Publishers and journals were identified from Beall's list on November 5, 2015. Email inquiries were sent to publishers of biomedical journals asking if they would add a co-author's name to any manuscript they had received for publication. After the initial inquiries that were sent to the first 200 publishers in alphabetical order demonstrated unethical responses, 63 additional emails were sent to randomly selected publishers. We randomly selected 75 stand-alone journals and sent emails to 64 (11 site links were not working). Email responses were categorized as ethical or unethical based on the journal's willingness to add coauthor names; responses generally not expected from a legitimate journal were considered unethical.

Results Of 906 publishers on Beall's list, 706 (77.9%) were screened (reasons for exclusion: nonworking links [n=184] and duplicates [n=16]), 400 (56.7%) of which published 4924 biomedical journals. Many publishers were located in India (n=119, 29.8%) and the United States (n=94, 23.5%). Among 835 stand-alone journals, 152 (18.2%) were biomedical journals. The overall response rate to the email inquiries was 44.5% (n=117) and 54.7% (n=35) for publishers and stand-alone journals, respectively. Nineteen publishers and 3 stand-alone journals agreed to add a coauthor name to an article they received without any specific contribution (**Table 22**). Forty-four publishers and 9 stand-alone journals "declined to add as coauthor." Overall, 63 publisher responses were unethical, 39 were ethical, and 15 were neutral; of the stand-alone journals, 17 were unethical, 11 were ethical, and 7 were neutral.

Conclusions Half or more of predatory publishers and stand-alone journals agreed to add a coauthor name without specification of any contribution, violating publication ethics and practice.

¹Scientific Writing and Epidemiology, Tata Consultancy Services, Thane, India, pravinbolshete@gmail.com

Conflict of Interest Disclosures: Dr Bolshete is an employee of Tata Consultancy Services, Thane, Mumbai. The views presented in this abstract are his and not those of his current or previous employers.

MONDAY

A Survey of Awareness of Authorship Criteria by Clinical Investigators and Medical Writers in China

Jing-ling Bao,¹ Xiu-yuan Hao,¹ Wei-zhu Liu,¹ Pei-fang Wei,¹ Yang Pan,¹ Jun-min Wei,¹ Young-mao Jiang¹

Objective To estimate the awareness of Chinese clinical investigators and medical writers about the authorship criteria defined by the International Committee of Medical Journal Editors (ICMJE) and the association between this awareness and the number of articles that are published.

Table 22. Summary of Responses Received from Predatory Publishers and Journals

Categories	Publishers, No. (%) (N=117)	Journals, No. (%) (N=35)
Declined to add as coauthor	44 (37.6)	9 (25.7)
No clear response	21 (17.9)	7 (20.0)
Agreed to add as coauthor	19 (16.2)	3 (8.6)
We will write article and publish	10 (8.5)	4 (11.4)
Agreed to write article but no clarity on publication	9 (7.7)	NA
Positive (but not clearly stated as yes)	4 (3.4)	2 (5.7)
Write yourself	4 (3.4)	2 (5.7)
Write yourself, we will publish	4 (3.4)	1 (2.9)
We will write, but cannot guarantee publication	1 (0.9)	NA
Other	1 (0.9)	NA
We will help you in writing	NA	7 (20.0)

Abbreviation: NA, not applicable.

Design A total of 935 clinical investigators and medical writers in China were surveyed online using a questionnaire with 4 questions. The first question addressed understanding of ICMJE authorship criteria in a multiple-choice format with 5 options: the first 4 options were the 4 specific authorship criteria defined by ICMJE, and the fifth option was that all those designated as authors should meet all 4 criteria for authorship. The second question asked about who should make the final decision regarding authorship and had 3 answer options: first author, corresponding author, and all authors. The third question asked whether the respondent had included someone who did not contribute to the work in an article. The fourth question asked about the number of papers submitted and published in Chinese and English-language journals. A χ^2 test was performed to analyze whether there was an association between the number of articles published and the awareness of authorship criteria.

Results Among the 935 by clinical investigators and medical writers from throughout mainland China invited to participate, 737 (78.8%) responded to the questionnaire. Of these, 726 (98.5%) provided usable responses. Regarding the first question on authorship criteria, 213 respondents (29.3%) gave the correct answer (the fifth option) and an additional 174 respondents (24.0%) indicated all first 4 options. Combined, 387 (53.3%) provided correct answers confirming an understanding of ICMJE criteria for authorship. Of the 339 respondents (46.7%) who did not select answers that demonstrated an understanding of the complete authorship criteria, 322 (95.0%), 254 (75.0%), 5 (1.5%), and 219 (64.6%) respondents chose the first to fourth options, respectively. There were no significant differences between the numbers of authors who published 0 to 2 articles and 3 or more articles in Chinese vs English-language journals. Only 100 respondents (13.8%) chose the option that all the authors should decide the order of authors, and there was no difference between the 2 groups. A total of 370 respondents (51%) indicated that they included someone who did not contribute to their work as an

Table 23. Comparison of Survey Responses by Different Published Groups

Options	Published in Chinese, No. (%)		P Value	Published in English, No. (%)		P Value
	0-2 Articles (N=195)	≥3 Articles (N=531)		0-2 Articles (N=462)	≥3 Articles (N=264)	
Answered the first question with the fifth ICMJE option	57 (29.2)	156 (29.4)	1.00	133 (28.8)	80 (30.3)	.67
Responded that all authors should choose the order of authorship	35 (17.9)	65 (12.2)	.05	66 (14.3)	34 (12.9)	.66
Included an honorary author	82 (42.1)	288 (54.2)	.004	234 (50.6)	136 (51.5)	.88

Abbreviation: ICMJE, International Committee of Medical Journal Editors.

author (honorary author); this was reported more frequently by authors who published 3 or more articles vs 0 to 2 articles in Chinese journals (54.2% vs 42.1%, $P = 0.004$) (**Table 23**).

Conclusions Many Chinese clinical investigators and medical writers are unaware of the ICMJE authorship criteria, and this unawareness has no association with the number of articles that are published. The existence of honorary authors is common. Editors should take responsibility for disseminating information regarding authorship criteria to authors.

¹Editorial Office, *Chronic Diseases and Translational Medicine*, Chinese Medical Association, 42 Dongsi Xidajie, Beijing, baojingling@cma.org.cn

Conflict of Interest Disclosures: None reported.

Funding/Support: None reported.

MONDAY

Survey of Authors' Views on Barriers to Preparation of Biomedical Research Manuscripts

June Oshiro,¹ Suzanne L. Caubet,² Kelly Viola,³ Jill M. Huber⁴

Objective A considerable body of biomedical research is presented as conference abstracts but never published as full manuscripts. We surveyed researchers at an academic medical institution, with the goals of (1) characterizing their self-identified barriers to manuscript preparation and (2) determining whether these barriers changed with increasing publication experience.

Design Select physicians and basic researchers (those attending noncompulsory workshops on publishing research) were surveyed with an emailed questionnaire from April 2009 through November 2015. All were employees of a single tertiary medical institution in the upper Midwestern United States. We asked them to report the number of published papers (coauthored) in the past 5 years and to indicate what was most difficult about preparing a manuscript for publication. Lack of time was presumed a major barrier; we thus asked them to list factors other than time. Two investigators performed a content analysis of deidentified free-text responses. The mean (SD) agreement between coders was 98% (2%), and the mean (SD) Scott π coefficient for interrater reliability was 0.81 (0.26). Participants' responses were stratified by publishing experience level (low [0-4 papers published in the past 5 years], medium [5-10 papers], and high [>10 papers]).

Results Of the 294 workshop participants, 201 (68.4%) responded to the survey. Of these, 77 (38.3%) had low experience, 53 (26.4%) had medium experience, and 71 (35.3%) had high experience in publishing. A total of 114 respondents (56.7%) listed multiple barriers to manuscript preparation (370 barrier items reported). Ten respondents (5.0%) did not indicate any barrier. The most frequently mentioned barriers, stratified by experience level, were not significantly different across groups (**Table 24**). Although the most common concerns overall (eg, organization and wording) appeared to be unaffected by author experience level, inexperienced researchers cited difficulty with defining the scope of the paper more often than highly experienced researchers, and with increasing experience, researchers had greater concerns about responding to reviewers and ensuring high-quality data presentation.

Conclusions The most commonly perceived barriers to manuscript preparation were not ameliorated by an increasing level of experience in publishing. Inexperienced researchers may benefit from mentoring to overcome difficulties with manuscript development. Most researchers, regardless of experience level, may be more productive if they had the assistance of a professional writer (not a ghostwriter) or editor who can help them draft the preliminary manuscript and follow the project through submission to ensure compliance with the journal-specific format and publication-quality figures and tables.

¹Section of Scientific Publications, Mayo Clinic, Rochester, MN, USA, oshiro.june@mayo.edu; ²Talent and Succession Management, Mayo Clinic, Rochester, MN, USA; ³Section of Scientific Publications, Mayo Clinic, Jacksonville, FL, USA; ⁴Division of Primary Care Internal Medicine, Mayo Clinic, Rochester, MN, USA

Conflict of Interest Disclosures: None reported.

MONDAY

Researchers' Awareness and Use of Authorship Guidelines: An International Survey

Sara Schroter,¹ Ilaria Montagni,² Elizabeth Loder,¹ Matthias Eikermann,³ Elke Schaeffner,⁴ Tobias Kurth¹

Objective To understand the degree to which international researchers are currently aware of and apply authorship guidelines in practice and to assess their perceptions of the fairness of authorship decisions.

Design In September 2016, we invited 12,646 corresponding authors of research papers submitted in 2014 to any of 18

Table 24. Most Common Barriers to Manuscript Preparation, Stratified by Publishing Experience

Barrier ^a	Experience Level of Respondents, No. (%) ^b			P Value ^c
	Low (n=77)	Medium (n=53)	High (n=71)	
Organizing information appropriately	19 (24.7)	16 (30.2)	16 (22.5)	.85
Succinct, effective wording	10 (13.0)	8 (15.1)	12 (16.9)	.64
Compliance with journal-specific format	9 (11.7)	9 (17.0)	10 (14.1)	.81
Defining scope of the paper	13 (16.9)	9 (17.0)	5 (7.0)	.08
Responding to reviewer concerns	2 (2.6)	5 (9.4)	7 (9.9)	.09
Creating publication-quality figures and tables	0 (0)	2 (3.8)	8 (11.3)	.002

^aTen respondents did not report any barriers (3 with a low level, 3 with a medium level, and 4 with a high level of experience).

^bA low level of experience was defined as 0 to 4 manuscripts published in the past 5 years; medium, 5 to 10 manuscripts; and high, more than 10 manuscripts.

^cDetermined by use of the Fisher exact test; comparison of low- and high-experience groups only.

BMJ journals in a range of specialties to complete an online survey.

Results We received 3859 responses (31%). Respondents varied in terms of research experience and worked in 93 countries. Of the 3859 respondents, 1326 (34%) reported that their institution had an authorship policy providing criteria for authorship; 2871 (74%) were “very familiar” with the International Committee of Medical Journal Editors (ICMJE) authorship criteria; and 3358 (87%) reported that these criteria were beneficial when preparing manuscripts. Furthermore, 2609 (68%) reported that their use was “sometimes” or “frequently” encouraged in their research setting. However, 2859 respondents (74%) reported that they had been involved in a study at least once where someone was added as an author who had not contributed substantially (honorary authorship), and 1305 (34%) were involved in a

study at least once where someone was not listed as an author but had contributed substantially to the study or article (ghost authorship). Only 740 (19%) reported that they had never experienced either honorary or ghost authorship; 1115 (29%) reported that they had experienced both at least once in their careers. **Table 25** lists respondents’ current perceived institutional support regarding authorship guidelines and their experience of authorship misappropriation by continent. While there are some differences by continent, there is no clear pattern. In regard to the last article that respondents coauthored, 2187 (57%) reported that explicit authorship criteria had been used to decide who should be an author, and 3088 (80%) felt the decision made was fair. When institutions frequently encouraged the use of authorship guidelines, authorship eligibility was more likely to be discussed at an early stage (817 of 1410, 58%) and was perceived as fairer (1273 of 1410, 90%) compared with infrequent encouragement (974 of 2449, 40% and 1891 of 2449, 74%).

Conclusions These results reflect current practice in international research across a range of specialties. While 74% of these authors are aware of guidelines, guidelines are not as frequently endorsed by institutions. Explicit encouragement of use of authorship criteria by institutions resulted in more favorable use of guidelines by author teams.

¹The *BMJ*, London, UK, sschroter@bmj.com; ²University of Bordeaux, Bordeaux, France; ³Critical Care Division, Massachusetts General Hospital, Boston, MA, USA; ⁴Charité–Universitätsmedizin Berlin, Berlin, Germany

Conflict of Interest Disclosures: Sara Schroter is a full-time employee at *BMJ*. Tobias Kurth is a consulting editor for *BMJ*. Elizabeth Loder received salary from *BMJ* for services as head of research, paid to her employer, the Brigham and Women’s Physician Organization.

Table 25. Experience of Authorship Misappropriation and Institutional Support by Continent of Respondent Author’s Main Institution^a

Characteristic	Respondents, No. (%)						
	All, N=3859	Africa, n=79	Asia, n=652	Europe, n=2073	North America, n=594	South America, n=90	Oceania, n=243
Institutional support							
Respondent’s institution has an authorship policy	1326 (34)	29 (37)	271 (42)	632 (31)	199 (34)	22 (24)	125 (51)
Respondent “very familiar” with ICMJE criteria	2871 (74)	56 (71)	429 (66)	1588 (77)	468 (79)	76 (84)	185 (76)
Use of explicit authorship guidelines “frequently” encouraged by respondent’s institution	1410 (37)	40 (51)	257 (39)	716 (35)	231 (39)	37 (41)	100 (41)
Authorship misappropriation							
Never experienced honorary authorship	929 (24)	22 (28)	219 (34)	449 (22)	129 (22)	16 (18)	67 (28)
Experienced honorary authorship at least once	2859 (74)	57 (72)	431 (66)	1621 (78)	461 (78)	74 (82)	176 (72)
Never experienced ghost authorship	2481 (64)	46 (58)	428 (66)	1368 (66)	371 (63)	63 (70)	166 (68)
Experienced ghost authorship at least once	1305 (34)	32 (41)	224 (34)	699 (34)	222 (37)	27 (30)	76 (31)
Experienced both honorary and ghost authorship	1115 (29)	28 (35)	180 (28)	605 (29)	192 (32)	24 (27)	65 (27)
Experienced neither honorary nor ghost authorship	740 (19)	17 (22)	176 (27)	355 (17)	100 (17)	13 (14)	56 (23)

^aPercentages do not all sum to 100% owing to missing data.

International Survey of Researchers' Experiences With and Attitudes Toward Coauthorship in the Humanities and Social Sciences

Tiffany Drake,¹ Bruce Macfarlane,² Mark Robinson¹

Objective To assess current attitudes toward and experiences of journal article coauthorship by researchers in the humanities and social sciences (HSS).

Design An online survey was distributed in June 2016 to 9180 researchers comprising editors of Taylor & Francis HSS journals, non-Taylor & Francis HSS editors, and researchers who published in Taylor & Francis HSS journals between July and September 2014. The survey included 13 questions about authorship and training/guidance followed by a scenario section, which presented researchers with a hypothetical situation to respond to from the perspective of their primary role in the publishing process.

Results A total of 894 participants (10%) from 62 countries completed all or part of the survey. Response rates varied by geographic location: Africa and Middle East, 11%; Australasia, 13%; Europe, 10%; Latin America, 18%; South and Southeast Asia, 10%; and United States and Canada, 9%. Among the 3 groups (authors, reviewers, and editors), respondents differed by mean number of articles published, sex, and age (**Table 26**) A total of 542 respondents (74%) reported that the typical number of authors per article in their area was 2 or more, and 501 respondents (56%) believed the incidence of coauthorship had increased since the beginning of their research careers. The most common reason given for increase of coauthorship was "increasing competition and greater performance-based pressures" (70%). The highest-scoring responses about common problems associated with coauthorship were "order in which author names should be listed" (52%) and "determining who should receive authorship credit" (43%). Respondents indicated that the following were important for determining authorship: "being responsible for the conception and/or design of a project" (79%); "being responsible for the analysis and/or interpretation of data" (81%); and "drafting the paper or revising it critically for intellectual content" (69%). Fewer respondents (18%) agreed that "giving final approval of the version of the paper to be published" was important for determining authorship. Respondents reported a reality gap, with "being a senior ranked member of the research team submitting a paper" and "being the research supervisor of a doctoral student whose paper gets published" considered less important in an ideal world vs the real world. Only 183 respondents (25%) reported that guidance on authorship was included in the research ethics policy of their institution, and 132 (18%) reported having received training or guidance from their institution in respect to determining academic authorship.

Conclusions With article coauthorship increasingly common in HSS, a need exists to address the attendant problems of authorship attribution. The results of this survey

Table 26. Demographic Characteristics of Respondents

Characteristics	Authors	Authors and Reviewers	Authors, Reviewers, and Editors	Total
Please select the role (author, etc) that best describes you, No. (%) (n=737) ^a	52 (7)	243 (33)	442 (60)	737
No. of articles published, mean	9.5	23	72	
No. of manuscripts reviewed, mean	NA	24	164	
First year of editorship, median	NA	NA	2008	
Sex, No. (%) (n=716) ^a				
Male	22 (46)	130 (54)	251 (60)	403
Female	26 (54)	106 (44)	168 (39)	300
Prefer not to say	0	4 (1)	6 (1)	10
Other	0	2 (0.8)	1 (0.2)	3
Age, No. (%) (n=717) ^a				
20-29	6 (13)	7 (3)	2 (0.5)	15
30-39	16 (33)	88 (36)	43 (10)	147
40-49	17 (35)	79 (33)	95 (22)	191
50-59	8 (17)	34 (14)	138 (32)	180
60-69	1 (2)	26 (11)	105 (24)	132
≥70	0	8 (3)	44 (10)	52
Region, No. (%) (n=737) ^a				
Africa and Middle East	1 (2)	20 (8)	30 (7)	51
Australasia	6 (12)	18 (7)	52 (12)	76
Europe	19 (37)	85 (35)	164 (37)	268
Latin America	1 (2)	8 (3)	12 (3)	21
South and Southeast Asia	4 (8)	17 (7)	25 (6)	46
United States and Canada	16 (31)	94 (39)	145 (33)	255
Unknown	5 (10)	1 (1)	14 (3)	20

Abbreviation: NA, not applicable.

^aRespondents did not answer all questions.

raise questions about the role institutions and publishers could play in providing clear ethical guidance and training for researchers and editors in these areas.

¹Taylor & Francis Group, Abingdon, UK, mark.robinson@tandf.co.uk; ²University of Southampton, Southampton, UK

Conflict of Interest Disclosures: As the employees of a commercial publisher (Taylor & Francis Group), Tiffany Drake and Mark Robinson report a potential conflict of interest where findings of the survey relate to aspects of the publishing process.

Financial Ties and Discordance Between Results and Conclusions in Trials of Weight Loss and Physical Activity Apps

Veronica Yank,¹ Sanjhavi Agarwal,¹ Rhea Red,² Amy Lozano¹

Objective Some of the most highly marketed mobile applications (“apps”) focus on weight loss and physical activity. The US Food and Drug Administration has declined to regulate them. It is unknown whether app research studies that receive financial support from commercial entities are similar to studies without commercial support in the degree to which they protect against bias. We sought to determine whether app studies with commercial financial ties were more or less likely than others to have discordant results and conclusions or other characteristics pertinent to assessing bias.

Design We performed a retrospective cohort study of English-language randomized clinical trials of weight loss or physical activity apps published through October 2016. We

searched PubMed, the Cochrane Database, EMBASE, and Web of Science using standardized approaches to identify articles that met initial literature search criteria. Two reviewers blinded to author name, affiliation, financial support, and conflict of interest disclosures performed title, abstract, and full-text review (as necessary) to determine final study eligibility. For included studies, 2 blinded reviewers independently assessed direction of study results (favorable or not favorable toward study app), direction of conclusions, and whether trial registration information and 7 design elements that protect against bias (from Cochrane Risk of Bias tool) were present. Financial ties were extracted after other assessments were complete. Financial ties were categorized as commercial (eg, ties to an app company) vs noncommercial (eg, ties to government or a foundation). Results and conclusions were defined as discordant if they disagreed in direction (results, not favorable; conclusions, favorable). Comparisons using Fisher exact tests and *t*-tests were performed using Stata statistical software (version 14.2; StataCorp).

Results Among 876 unique articles identified on initial literature search, 17 met inclusion criteria: 7 (41%) with commercial ties and 10 (59%) with noncommercial ties (Table 27). A smaller percentage of commercial trials

Table 27. Characteristics of Randomized Clinical Trials of Weight Loss and Physical Activity Apps According to Commercial and Noncommercial Financial Ties*

Financial Ties, Study No.	Favorable Results	Favorable Conclusions	Discordant Findings	Trial Registration	Protections Against Bias
Commercial (n=7)					
1	NF	F	Yes	No	0
2	NF	F	Yes	No	0
3	NF	F	Yes	No	2
4	NF	NF	No	Yes	2
5	NF	F	Yes	No	0
6	F	F	No	No	2
7	NF	F	Yes	No	1
Trials, No. (%)	1 (14)	6 (86)	5 (71)	1 (14)	Mean (SD), 1 (0.4)
Noncommercial (n=10)					
1	F	F	No	Yes	3
2	NF	F	Yes	Yes	4
3	NF	NF	No	Yes	5
4	F	F	No	Yes	2
5	F	F	No	Yes	4
6	F	F	No	Yes	4
7	F	F	No	Yes	3
8	NF	NF	No	Yes	5
9	F	F	No	Yes	6
10	NF	NF	No	Yes	4
Trials, No. (%)	6 (60)	7 (70)	1 (10)	10 (100)	Mean (SD), 4 (0.2)

Abbreviations: F, favorable toward study app; NF, not favorable.
*Trial registration indicates that the study reports being registered in a national or international trial registry (eg, clinicaltrials.gov). Protections against bias was scored (range, 0-7) using the Cochrane Risk of Bias tool, with higher scores indicating a higher number of study design elements that protect against bias. Design elements assessed include random sequence generation, allocation concealment, blinding of participants and/or personnel, blinding of outcome assessment, completeness of outcome data, avoidance of selective reporting, and avoidance of other sources of bias.

reported favorable results than did noncommercial trials (14% vs 60%, respectively; $P = .13$), whereas a larger percentage reported favorable conclusions (86% vs. 70%; $P = .60$). As a result, commercial trials had significantly greater discordance between results and conclusions than noncommercial trials (71% vs. 10%; $P = .04$). They also were less likely to report trial registration (14% vs 100%; $P = .001$) and had fewer design elements that protect against bias (mean [SD], 1 [0.4] vs 4 [1.2]; $P < .001$).

Conclusions Randomized clinical trials of weight loss and physical activity apps with commercial financial ties were more likely to have discordance between results and conclusions and less likely to meet registration and design standards.

¹University of California, San Francisco, San Francisco, CA, USA, veronica.yank@ucsf.edu; ²San Francisco State University, San Francisco, CA, USA

Conflict of Interest Disclosures: None reported.

Funding/Support: Dr Yank is supported by a career development award from the National Institute of Diabetes and Digestive and Kidney Diseases (K23DK097308).

Role of Funder/Sponsor: The National Institute of Diabetes and Digestive and Kidney Diseases had no role in the design, performance, or interpretation of findings of this study.

TUESDAY

Bias Arising From the Use of Patient-Reported Outcome Measures

Joel J. Gagnier,^{1,2} Jianyu Lai,¹ Chris Robbins¹

Objective The objective was to assess the bias in outcomes effects associated with the use of patient-reported outcome (PRO) measures of varying psychometric quality in peer-reviewed clinical studies.

Design A literature search was conducted using PubMed (January 1, 2011, to December 31, 2016) to search for randomized and observational studies (cohort, case-control studies) published in the top 5 orthopedic journals (as ranked by their Impact Factors), in humans that used PRO measures, for rotator cuff disease. The study design, sample size, standard aspects of risk of bias (high or low using Cochrane Methods, such as randomization methods, allocation concealment, and blinding) for each study type, types of PRO outcome measures used, data related to the PRO measure results in all groups, measures of effect (odds ratio, relative risk, incidence risk ratio, absolute risk reduction, mean change, etc), and confidence intervals or other measures of variance were extracted. Continuous measures of effect were transformed by dividing the effect estimate by the standard deviation. PRO measures were given numerical ratings of quality based on a systematic review of evidence for their psychometric properties. Linear regression analyses were performed to determine whether PRO measure quality was associated with the magnitude of effects and the influence of a variety of covariates on this relationship.

Results Overall, 162 articles were included for this study across 5 high-impact orthopedic journals. Of the studies, 80% were observational and 20% randomized. Studies included 1 to 7 PRO measures. Greater than 75% of the included studies did not justify the use of PRO and greater than 50% did not describe the PRO details. Linear regression revealed that lower-quality PROs had larger estimates of effect, and by contrast, higher-quality PROs had smaller estimates of effect ($n = 123$; $\beta = -0.21$, 95% CI, -0.43 to -0.02 ; $P = .03$). In univariable regression analyses, we also found that a longer follow-up period (range, 0.1 to 96 months) predicted slightly increased effect estimates ($n = 96$; $\beta = 0.05$; 95% CI, 0.02 to 0.09 ; $P = .002$).

Conclusions PRO measures with poor or unknown psychometric properties bias (ie, inflate) the estimates of treatment effect in clinical research of rotator cuff disease. To our knowledge, this is the first empirical evidence, to date, that variations in the quality of PRO measures bias treatment effect estimates. Researchers and clinicians using data from PROs must be cautious to explore the quality of that measure so as to not mislead decision making from biased outcomes.

¹Department of Orthopaedic Surgery, University of Michigan, Ann Arbor, MI, USA, jgagnier@med.umich.edu; ²Department of Epidemiology, University of Michigan, Ann Arbor, MI, USA

Conflict of Interest Disclosures: None reported.

MONDAY

Discrepancies in Risk of Bias Judgments for Randomized Trials of Acupuncture Included in More Than 1 Cochrane Review

Yonggang Zhang,^{1,2} Linli Zheng,² Youping Li,^{1,2} Mike Clarke,^{1,3} and Liang Du^{1,2}

Objective To assess consistency in risk of bias judgments for randomized trials of acupuncture included in more than 1 Cochrane Review.

Design We identified randomized trials of acupuncture that appeared in more than 1 Cochrane Review and retrieved all risk of bias judgments for these trials. We assessed the consistency of judgments (high risk of bias, low risk of bias, and uncertain) for the 5 domains in the Cochrane risk of bias tool: random sequence generation, allocation concealment, blinding, incomplete outcome data, and selective reporting. Reviews that did not report all 5 domains were included in the analyses of the domains they did report.

Results We identified 90 Cochrane Reviews that included at least 1 randomized trial of acupuncture, comprising a total of 1692 trials. After checking the reviews, 31 trials were identified in more than 1 review (in a total of 28 Cochrane Reviews). Thirty trials appeared in 2 reviews and 1 trial appeared in 3 reviews. For all 31 trials, we found a total of 121 judgments for the 5 domains. Overall, 50% (60 of 121) of these judgments were different (**Table 28**), with most of these differences being the categorization as uncertain in 1

Table 28. Judgments of the Risk of Bias for 31 Randomized Trials of Acupuncture

Risk of Bias Domain	Total No. of Trials	Judgments Agreed, No. (%)	Judgments Disagreed, No. (%)	Disagreements		
				Unclear/Low Risk	High Risk/Unclear	High Risk/Low Risk
Random sequence generation	22	15 (68)	7 (32)	7	0	0
Allocation concealment	29	15 (52)	14 (48)	10	3	1
Blinding ^a	24	8 (27)	16 (63)	5	10	4
Incomplete outcome data	24	15 (63)	9 (27)	3	1	5
Selective reporting	22	8 (36)	14 (64)	11	3	0

^aBlinding was assessed by "blinding treatment provider" and "blinding outcome assessor" in 3 studies, disagreements are in both 2 subdomains; therefore, the total number of disagreements in the domain of blinding is 19.

review but high or low risk in another. Relatively good agreement was found for random sequence generation (68%) and incomplete outcome data (63%). Five of the 9 discrepant trials with incomplete outcome data had the most extreme inconsistency (ie, high risk of bias in 1 review but low risk of bias in the other review). Agreement was 52% for allocation concealment but only 27% for blinding and 36% for selective reporting.

Conclusions Use of acupuncture as example of the assessment for bias in Cochrane trials may be a limitation of this study given the concerns about blinding in trials of acupuncture. However, this analysis shows that there are large discrepancies in risk of bias judgments between Cochrane Reviews that assessed the same acupuncture randomized trial, which may cast doubt on the much commoner situation, when a trial is assessed once only. Further work is needed to improve the application of the Cochrane risk of bias tool. The collation of judgements for all randomized trials in a central, standardized database of risk of bias may be helpful.

¹*Journal of Evidence-Based Medicine*, m.clarke@qub.ac.uk;

²Chinese Cochrane Centre, West China Hospital of Sichuan University, Chengdu, China; ³Centre for Public Health, Queen's University Belfast, Belfast, UK

Conflict of Interest Disclosures: Dr Li is a member of the Peer Review Congress Advisory Board but was not involved in the review or decision of this abstract.

Funding: None reported.

MONDAY

Gender Bias in Funding of Proposals Submitted to the Swiss National Science Foundation

François Delavy,¹ Anne Jorstad,¹ Matthias Egger¹

Objective The Swiss National Science Foundation is the largest public research funder in Switzerland. Its Research Council assesses proposals and makes funding decisions. Proposals can be subject to budget cuts. Our objective was to investigate whether success rates and cuts in budgets differed between male and female applicants.

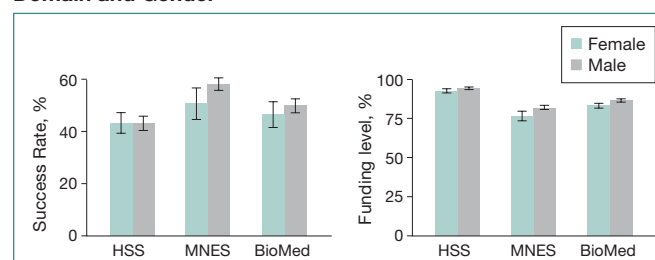
Design We analyzed the decisions for proposals submitted from 2014 to 2016. We computed average success rates and funding levels (yearly approved amount of funding divided by

yearly requested amount of funding) by gender and domain of research: Humanities and Social Sciences (HSS), Mathematics, Natural and Engineering Sciences (MNES); and Biology and Medicine (BioMed). Funded projects received a mean annual support of 140,000 CHF (~144,000 USD). We performed 2 regression analyses to adjust for potential confounders, including submission year and characteristics of the applicants (affiliation, discipline, age, nationality, first-time applicants): a logistic regression for success on all proposals (n=5687) and a linear regression for funding level on approved proposals (n=2824).

Results The average success rate of female applicants was similar to male applicants in HSS, but 7.8% lower in MNES and 3.5% lower in BioMed (**Figure 8**). In the multivariable logistic regression model the odds of approval continued to be lower for proposals submitted by women in MNES (odds ratio compared with men, 0.72 [95% CI, 0.54-0.96]); but not for HSS (95% CI, 0.77-1.15) and BioMed (95% CI, 0.77-1.25). Furthermore, the budgets of proposals from female applicants were cut more substantially than the budgets of male applicants. Funding levels were 2.7% lower in HSS, 5.7% lower in MNES, and 3.2% lower in BioMed in female compared with male applicants. In the multivariable linear regression model, the funding level was 2.5% (95% CI, 0.3%-4.6%), 6.0% (95% CI, 2.3%-9.7%), and 3.6% (95% CI, 0.9%-6.0%) lower for female compared with male applicants in domains HSS, MNES, and BioMed, respectively.

Conclusions Proposals from female applicants had lower success rates in MNES, BioMed, and differences persisted in models adjusted for potential confounding factors in MNES. Furthermore, budget cuts were more substantial for female than for male applicants in all domains, and differences again

Figure 8. Success Rate and Funding Level by Research Domain and Gender



HSS Indicates humanities and social sciences; MNES, mathematics, natural and engineering sciences; BioMed, biology and medicine.

persisted in adjusted analyses. These results are compatible with a gender bias but cannot prove the existence of such bias.

¹Swiss National Science Foundation, Bern, Switzerland, francois.delavy@snf.ch

Conflict of Interest Disclosures: None reported.

TUESDAY

Prevalence of High or Unclear Risk of Bias Assessments in Diagnostic Accuracy Studies Included in Cochrane Reviews

Nicola Di Girolamo,¹ Reint Meursinge Reynnders,² Alexandra Winter^{3,4}

Objective The validity of systematic reviews of diagnostic accuracy is dependent on the extent of bias in included primary studies. The objective of this study was to assess the risk of bias in primary studies of diagnostic accuracy included in Cochrane reviews according to the Quality Assessment for Studies of Diagnostic Accuracy 2 (QUADAS-2) tool.

Design All systematic reviews of diagnostic accuracy published in the Cochrane database in 2016 that used the QUADAS-2 quality assessment tool and that reported results with the risk of bias figure were eligible. The primary outcome was the prevalence of high or unclear risk-of-bias scores for the 4 QUADAS-2 domains—patient selection, index test, reference standard, and flow and timing—among the primary studies in these diagnostic accuracy reviews. Two investigators selected the eligible reviews and assessed the risk of bias scores for primary studies. Disagreements were resolved by consensus.

Results Of 46 eligible systematic reviews, 35 were included, 6 were excluded because they did not use the QUADAS-2 tool, 4 were excluded because they lacked the risk-of-bias figure, and 1 review was a duplicate. A total of 1045 primary studies with 4133 bias assessments were identified; 1044 reported risk of bias for patient selection, 1002 for index test, 1044 for reference standard, and 1043 for flow and timing. A total of 2319 of 4133 domains (56%) were assessed as having high or unclear risk of bias, with 1814 (44%) having low risk of bias. For all domains except flow and timing, the majority of outcomes were scored as having high or unclear risk of bias.

Conclusions Primary studies in systematic reviews of diagnostic accuracy are often rated as having high or unclear risk of bias by QUADAS-2 criteria. Inclusion of such studies in systematic reviews and meta-analyses may jeopardize final results and interpretation. This study is limited by the lack of accounting for clustering of risk-of-bias assessments within primary studies evaluated in systematic reviews and for the exclusion of non-Cochrane reviews, which may limit the generalizability of the findings. Although further investigation is indicated to evaluate whether additional training of systematic reviewers would decrease the prevalence of high and unclear risk of bias, the findings point to a need to

improve the conduct and reporting of diagnostic accuracy studies.

¹EBMVet, Cremona, Italy, nicoladiggi@gmail.com; ²Department of Oral and Maxillofacial Surgery, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands; ³*Journal of the American Veterinary Medical Association*, Schaumburg, IL, USA; ⁴American Veterinary Medical Association, Schaumburg, IL, USA

Conflict of Interest Disclosures: None reported.

Funding/Support: This study was self-funded.

TUESDAY

Assessment of Agreement Between Reviewers in the Open Postpublication Peer Review Process of F1000Research

Tiago Barros,¹ Liz Allen¹

Objective F1000Research operates an author-driven, open, and postpublication peer review model. The identity of the reviewer and the peer review report including its recommendation are made public immediately after submission by the reviewer. This study aimed to identify any potential influence of the first published peer review on the recommendation of the second reviewer, as measured by the agreement between the 2 recommendations and the time between them.

Design Bringing together a dataset of articles published between July 2012 and February 2017 and associated open peer review reports, we analyzed the agreement among reviewers depending on the time between reports. Only articles presenting original research or methods were included. Articles where the time gap between the 2 reviewer reports was longer than a year (365 days) were excluded. The recommendations (“approved,” “approved with reservations,” or “not approved”) of the first 2 reviewers were recorded, as well as the published date of the reports. Cohen κ was used to measure interrater reliability, and its change with time between reports was used to assess potential bias. In the absence of survey data on whether the second reviewer had read a previous report before submitting their recommendation, reports published within the same day are considered the control group.

Results The analyzed dataset contained 1,133 articles and 2,266 reviewer reports, ie the first 2 reviewer reports of each article. The median (interquartile range) time between the first 2 peer reviews was 18 (6–52) days. In aggregate, the breakdown of the peer review decision (“approved,” “approved with reservations,” or “not approved”) across the dataset was virtually identical between the 2 reviewers (724 [63.9%], 355 [31.3%], and 54 decisions [4.8%] vs 705 [62.2%], 372 [32.8%], and 56 [4.9%] decisions, respectively). However, comparing the recommendations made for each article individually, the Cohen κ was 0.330 (compared with 0.282 for the control group), indicating only a fair agreement between the reviewers. Moreover, the Cohen κ changed

minimally with the length of time between the peer review publication dates (**Table 29**).

Conclusions Our analysis of the *F1000Research* open peer reviews found that the agreement between reviewers did not change substantially with the time gap between peer reviews. The second reviewer does not seem to be systematically influenced by the ability to see the recommendation of an earlier reviewer. This is an important finding and something to continue to monitor as the momentum and acceptance of open peer review models, and open science more broadly, continues to grow.

¹*F1000*, London, UK, tiago.barros@f1000.com

Conflict of Interest Disclosures: Dr Barros is the Product Strategy Manager of *F1000*, and Dr Allen is the Director of Strategic Initiatives of *F1000*.

MONDAY

An Update on Reporting Bias in the Antidepressant Literature: An FDA-Controlled Examination of Drug Efficacy

Erick H. Turner,^{1,2,3} Sepideh Alavi,² Andrea Cipriani,⁴ Toshi Furukawa,⁵ Ilya Ivlev,¹ Ryan McKenna,³ Yusuke Ogawa⁵

Objective We previously investigated the influence of reporting bias on the apparent proportion of statistically significant trials and effect size estimates for antidepressant medications approved through 2004. We update those findings here for medications approved since 2004.

Design We identified antidepressants approved by the US Food and Drug Administration (FDA) since 2004. We downloaded corresponding medical and statistical reviews from *Drugs@FDA*, identified phase 2 and 3 double-blind placebo-controlled efficacy trials, extracted summary statistics on each trial's primary outcome, and extracted the FDA's judgment as to whether each trial provided evidence of efficacy (statistical superiority to placebo on the primary outcome). For each FDA-registered trial, we searched the published literature for corresponding journal publications, extracted from the results sections summary statistics on the effect size for the stated primary outcome and whether the publication conveyed that the drug was effective, and compared trial outcome data from the FDA vs journal publications. We conducted 2 meta-analyses using the

published literature and using FDA data and compared the resulting effect size (standardized mean difference (SMD) values using meta-regression. We repeated the meta-analysis comparison combining newer- and older-cohort datasets. We contrasted the extent of effect size inflation (bias) in the old vs new cohorts.

Results Four antidepressant drugs were approved by the FDA since 2004: desvenlafaxine, levomilnacipran, vilazodone, and vortioxetine. As with older antidepressants, 50% of the FDA trials (n = 15) evaluating the newer drugs showed a statistically significant difference (**Table 30**). Reporting bias inflated the proportion of apparently positive trials (+21%), but less compared with the older cohort (+43%). Within the nonsignificant trials, the percentage published transparently (trial published and in agreement with FDA) increased significantly from 8.3% (older drugs) to 40% (newer drugs) (P = .01). Nevertheless, when nonsignificant and significant trials were combined, the rate of transparent publication was significantly greater for significant compared with nonsignificant trials (P = 5 × 10⁻¹⁹). In meta-analyses, the boost in SMD due to reporting bias diminished from 0.10 (older drugs) to 0.05 (newer drugs). Differences between FDA- and journal-based effect size values using meta-regression were statistically significant for the older drugs (P = .001), not statistically significant for newer drugs (P = .25), but statistically significant when older and newer drugs were combined (P = .003).

Conclusions Reporting bias continues in the antidepressant clinical trial literature but findings with newer drugs compared with older drugs suggest a decrease in magnitude of reporting bias due to more transparent disclosure of nonstatistically significant clinical trial results.

¹Oregon Health & Science University, Portland, OR, USA, eturnermd@gmail.com; ²VA Portland Health Care System, Portland, OR, USA; ³Scientific Resource Center, Agency for Health Research Quality, Portland, OR, USA; ⁴University of Oxford, Oxford, UK; ⁵Kyoto University Graduate School of Medicine, Kyoto, Japan

Conflict of Interest Disclosures: Erick H. Turner receives protected time for research through Kyoto University Graduate School of Medicine. Andrea Cipriani has served as an expert witness for a patent litigation case about quetiapine extended release. Toshi Furukawa has received lecture fees from Eli Lilly, Janssen, Meiji, MSD, Otsuka, Pfizer and Tanabe-Mitsubishi and consultancy fees from Sekisui Chemicals; he has received royalties from Igaku-Shoin and Nihon Bunka Kagaku-sha publishers; and he has received research support from Mochida and Tanabe-Mitsubishi.

MONDAY

Prevalence of Comparative Effectiveness Trials of Surgical vs Medical Interventions

Anaïs Rameau,¹ Anirudh Saraswathula,² Ewoud Schuit,³ John P. A. Ioannidis⁴

Objective Surgical and medical (drug) treatment options exist for many conditions. Practitioners from different specialties often perform or prescribe these interventions, and it is unknown how often the options are directly

Table 29. Cohen κ and Days Between Peer Reviews for 1133 Articles

Days between peer reviews	Articles, No.	Cohen κ
0 (published simultaneously)	54	0.282
1 to 5	205	0.316
6 to 12	207	0.352
13 to 27	218	0.340
28 to 64	229	0.312
65 to 365	220	0.302

Table 30. Reporting Bias: Relative Magnitude With Newer vs Older Antidepressants

	Older Cohort of Antidepressants	Newer Cohort of Antidepressants	New + Old Combined	Difference, New vs Old
No. of drugs	12	4	16	−8
No. of trials	74	30	104	−44
Proportion of statistically significant trials per data source				
FDA	51% (38/74)	50% (15/30)	51% (53/104)	−1% (≈No change)
Published literature	94% (48/51)	71% (17/24)	87% (65/75)	−23% (71% vs 94%)
Difference due to pub bias	+43%	+21%	36%	−22%
NS trials, % published	8.3% (3/36)	40% (6/15)	18% (9/51)	+9.7% (=.01 ^a)
All trials (significant + NS), rate of transparent publication vs outcome ^b	$P = 2.5 \times 10^{-16}$ (37, 3/1, 33)	$P = .0007$ (15, 6/0, 9)	$P = 5 \times 10^{-19}$ (52, 9/1, 42)	NA
Effect size^c				
FDA	0.31 (0.27, 0.35)	0.24 (0.18, 0.30)	0.28 (0.25, 0.32)	−0.07
Published literature	0.41 (0.36, 0.45)	0.29 (0.23, 0.36)	0.36 (0.32, 0.40)	−0.12
Effect size inflation, FDA vs published literature				
Arithmetic change	+0.10	+0.05	+0.08	−0.05
Meta-regression	$P = .001$	$P = .25$	$P = .003$	$P = .25$

Abbreviations: FDA, US Food and Drug Administration; NA, not available.

^aFisher exact test, (a) cohort old vs new, (b) published in agreement with the FDA (3, 6 / 33, 9).

^bFisher exact test on a 2×2 table with factors (trial significant vs published in agreement with the FDA).

^cStandardized mean difference (Hedges g [95% CI]).

compared in randomized clinical trials (RCTs). We aimed to investigate the prevalence of comparative effectiveness trials of surgical vs medical interventions by assessing Cochrane systematic reviews that address surgical interventions and recording whether they include RCTs comparing surgical interventions with medical interventions.

Methods We searched the Cochrane Library from inception until September 2015 to identify all published Cochrane meta-analyses of surgical interventions using the search term *surg** in “search all text.” Only meta-analyses presenting evidence from RCTs were eligible. Medical intervention was defined as the administration of medication. Interventional radiology, dental, gastroenterologic, and cardiologic procedures were excluded. For Cochrane reviews meeting the criteria, we established the number of surgical vs medical RCTs used in the meta-analysis. We determined the standardized mean difference in effect size between surgical and medical interventions within each meta-analysis to assess which interventions showed overall better results.

Results Of 3475 reviews identified in our search, 33 including 54 unique RCTs of surgical vs medical interventions met inclusion criteria; in another 46, the authors intended to compare surgical and medical interventions but no RCTs were found. Across the 33 meta-analyses, medical intervention was found superior to surgical intervention for the primary outcome in 5, and surgical intervention was found superior to medical intervention in 9. The remaining 19 reviews were inconclusive.

Conclusions We identified a lack of RCTs directly comparing medical and surgical interventions. More such trials are needed to document whether surgical and medical

interventions are equally good or one is better than the other in different conditions.

¹Department of Otolaryngology–Head and Neck Surgery, University of California-Davis, Sacramento, CA, USA, anaais@stanford.edu; ²Stanford University School of Medicine, Stanford, CA, USA; ³Julius Center, University Medical Center Utrecht, Utrecht, the Netherlands; ⁴Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA, USA

Conflict of Interest Disclosures: John P. A. Ioannidis is a Peer Review Congress Assistant Director but was not involved in the review of or decision for this abstract.

Funding/Support: Stanford Society of Physician Scholars. The funding agency had no role in the research.

TUESDAY

Frequency of Citation of Clinical Trials With a High Risk of Bias

Priyanka Desai,¹ Mary Butler,¹ Robert L. Kane¹

Objective To compare citation rates of high, medium, or low risk of bias (ROB) clinical trials included in systematic reviews.

Design We used a convenience sample of systematic reviews published between 2011 and 2014 by 4 journals, the Cochrane Collaboration, and the Agency for Healthcare Research and Quality Evidence-based Practice Centers to identify a sample of clinical trials. We abstracted the trial ROB assigned by review authors, or, for trials rated as having an unclear ROB, we assigned ROB ratings using the Cochrane ROB tool. We converted quality and numeric ratings to ROB ratings and assigned overall ratings to studies that were assessed with only domains of the Cochrane Tool. We used Google Scholar

Table 31. Mean Citations by Risk of Bias (ROB) and Correlations Between ROB and Citations

Intervention Category (N)	Mean Citations				Mean Impact Factor (Range)	Spearman Correlation Coefficient (ROB)
	Low ROB (N=500)	Medium ROB (N=628)	High ROB (N=328)	Overall		
Device (158)	40.15	95.18	26.77	62.56	3.60 (0.08-53.30)	0.24 ^a
Lifestyle (156)	105.05	113.69	41.81	95.20	6.27 (0.42-54.42)	0.19 ^a
Organizational (81)	83.57	59.67	47.75	62.04	4.38 (0.57-53.49)	0.12
Pharmaceutical (856)	125.47	76.67	92.31	98.48	5.86 (0.17-54.42)	0.19 ^a
Psychosocial (148)	127.65	71.51	78.07	86.65	3.66 (0.33-17.57)	0.11
Surgical (57)	167.24	28.18	19.00	67.72	6.43 (0.15-52.59)	0.13
All Categories (1456)	116.96	77.57	71.81	89.79	5.38 (0.08-54.42)	0.19 ^a

Abbreviation: ROB, risk of bias.

^a $P < 0.05$

as a source for citation rates and Journal Citation Reports as a source for citing the Impact Factors of journals. We reported mean trial citation rates by low, medium, and high ROB and calculated Spearman correlation coefficients to assess the association between ROB and citations and citations and impact factors. We reported findings for cited trials overall and stratified by the type of intervention .

Results Of 76 systematic reviews, 55 reported sufficient ROB information. Of 1456 trials with ROB ratings, 34.4% (500) were rated with low, 43.1% (628) with medium, and 22.5% (328) with high ROB. Across all intervention categories, low ROB studies were more frequently cited than high ROB studies (**Table 31**). The correlation between ROB and citation rates was generally weak but statistically significant for trials of device, lifestyle, and pharmaceutical interventions. Across all intervention categories, there was a moderate correlation between citation rates and Impact Factors (correlation coefficients from 0.46 to 0.62, $P < .01$)

Conclusions Weak correlations between clinical trial ROB and citation rates suggest that ROB is not a meaningful factor in the decision to cite a publication. We did not examine the context in which high ROB trials were cited; however, their inclusion in systematic reviews may pose a risk to evidence-based practice when ROB is not fully considered. High ROB ratings are difficult to interpret, and a variety of factors may contribute to the rating. The field should consider how to adequately inform readers and end users about the potential biases of published clinical trials.

¹University of Minnesota School of Public Health, Minneapolis, MN, USA, butloo92@umn.edu

Conflict of Interest Disclosures: None reported.

Funding/Support: None reported.

Acknowledgments: Dr Kane passed away on March 6, 2017.

Bibliometrics and Scientometrics

TUESDAY

The Clinical Impact of Published Trials

Ashwini R. Sehgal^{1,2,3}

Objective A key goal of medical journals is to influence clinical practice. However, there are no objective, reproducible, or comprehensive measures of the clinical impact of articles published in medical journals. UpToDate is an online, continuously updated information resource used by more than 1 million clinicians to obtain specific recommendations regarding diagnosis and treatment. Each UpToDate chapter is generally written by 1 to 3 physician authors who are also responsible for selecting articles to cite. A section editor, deputy editor, and peer reviewers provide additional input. We reasoned that citations in UpToDate may provide a useful measure of the clinical impact of published articles. In previous work, we ranked journals based on their citations in UpToDate and compared this type of ranking with journal impact factors. We found little relationship between journal rankings based on UpToDate citations and those based on impact factor, which indicates that these are distinct markers. We then sought to use UpToDate citations to assess the clinical impact of specific trials 2 to 10 years following publication.

Design We selected all 1527 clinical trials published in 3 general medical journals (the *New England Journal of Medicine*, *JAMA*, and the *Annals of Internal Medicine*) and 3 nephrology journals (the *Journal of the American Society of Nephrology*, *Kidney International*, and the *American Journal of Kidney Diseases*) in 2006, 2008, 2010, 2012, and 2014 and then determined how many of these published articles were cited in UpToDate in March 2016.

Results Of all 1527 clinical trials published in the 6 journals, 76% (95% CI, 74%-78%) were cited in UpToDate in March 2016. Trials published in the 3 medical journals were more likely than trials published in the 3 nephrology journals to be cited in UpToDate (85% vs 49%; difference, 36% [95% CI, 31%-42%]). From 2006 to 2014, the proportion of medical journal trials cited in UpToDate showed little variation (range, 84%-89%). From 2006 to 2014, the proportion of nephrology journal trials cited in UpToDate increased from 37% to 57% (difference, 20% [95% CI, 5%-33%]).

Conclusions A large proportion of clinical trials published in these 6 journals may have influenced clinical practice by informing specific recommendations that guide clinicians

regarding diagnosis and treatment. The proportion varies by journal type and year of publication. The limitations of this study include the small sample of journals and the lack of a gold standard of clinical impact.

¹Division of Nephrology, MetroHealth Medical Center, Cleveland, OH, USA, axs81@cwru.edu; ²Center for Reducing Health Disparities, Case Western Reserve University, Cleveland, OH, USA; ³*Annals of Internal Medicine*, Philadelphia, PA, USA

Conflict of Interest Disclosures: None reported.

MONDAY

Association Between the Journal Evaluation Program of the Korean Association of Medical Journal Editors (KAMJE) and Change in Quality of Member Journals

Hee-Jin Yang, MD, PhD^{1,2}; Se Jeong Oh, MD, PhD^{1,3}; Sung-Tae Hong, MD, PhD^{1,4}

Objective In 1997 the Korean Association of Medical Journal Editors (KAMJE) instituted a program to evaluate Korean medical journals. Journals were evaluated on criteria such as timeliness, quality of editorial work, and adherence of bibliography and citations to high standards. Journals that passed the initial evaluation process were indexed in KoreaMed, the Korean version of PubMed. Here, we report changes in measures of quality among the KoreaMed-indexed journals that were associated with the evaluation program after 7 years.

Design Quality measures used in the study comprised self-assessment by journal editors and assessment of the journals by KAMJE reviewers and by Korean health science librarians. Each used detailed criteria to score the journals on a scale of 0 to 5 or 6 in multiple dimensions. We compared scores at baseline evaluation and after 7 years for 129 journals and compared improvements in journals indexed vs. not-indexed by Web of Science.

Results Among 251 KAMJE member journals at the end of 2015, 227 passed evaluation criteria and 129 (56%) had both baseline and 7-year follow-up assessment data. The journals showed improvement overall (increase in mean [SD] score from baseline, 0.53 [0.48]; 95% CI, 0.44-0.61; $P < .001$) and within each category of evaluation (mean [SD] increase by editor's assessment, 0.14 [0.58]; 95% CI 0.04-0.26; $P = .007$; reviewer's, 0.43 [0.76]; 95% CI, 0.29-0.57; $P < .001$; and librarian's, 1.98 [1.15]; 95% CI, 1.77-2.18, $P < .001$) (**Table 32**). Before the foundation of KAMJE in 1996, there were only 5 Korean medical journals indexed in the MEDLINE and 1 indexed by Web of Science (SCI). By 2016, there were 24 journals listed in MEDLINE and 34 journals indexed in Web of Science (SCI). There was no statistically significant difference in scores on initial assessments between 21 SCI-indexed and 108 non-indexed journals, but the scores of the SCI-indexed journals were significantly higher on follow-up assessments (mean [SD], 3.99 [0.37] vs 3.38 [0.43]).

Table 32. Change of Scores of Journal Evaluations Between Initial and Follow-up Assessments (N=129)

Evaluator	Mean (SD)			95% CI	P value ^b
	Initial Scores ^a	Follow-up Scores ^a	Difference		
Editors	3.68 (0.41)	3.83 (0.48)	0.14 (0.58)	0.04-0.26	.007
Reviewers	3.28 (0.51)	3.71 (0.67)	0.43 (0.76)	0.29-0.57	<.001
Librarians	0.68 (0.56)	2.66 (1.11)	1.98 (1.15)	1.77-2.18	<.001
Total	2.93 (0.31)	3.46 (0.47)	0.53 (0.48)	0.44-0.61	<.001

^aScores for each are rated based on detailed description (KAMJE's criteria for journal evaluation); range, 0.09-5.09.

^bPaired t-test

Conclusions These results suggest an association between a program of assessment by editors, reviewers, and librarians and improvement in quality of KAJME member journals. The increase in the number of KAJME member journals indexed in international databases also suggests that the KAMJE program is successful at improving journal quality.

¹The Korean Association of Medical Journal Editors, Seoul, Korea, hst@snu.ac.kr; ²Department of Neurosurgery, Seoul National University Boramae Hospital, Seoul, Korea; ³Department of Surgery, The Catholic University of Korea, Incheon St. Mary's Hospital, Incheon, Korea; ⁴Department of Parasitology and Tropical Medicine, Seoul National University College of Medicine, Seoul, Korea

Conflict of Interest Disclosures: None reported.

Funding/Support: Financial support by a grant from the Korean Association of Medical Journal Editors (KAMJE), 2015.

Role of the Funder/Sponsor: The funder supplied data of the journal evaluation program but was not otherwise involved in the study.

MONDAY

Association of Publication Rate With the Award of Starting and Advanced Grants

David Pina,¹ Lana Barać,^{2,3} Ivan Buljan,³ Ana Marušić³

Objective To analyze the association of European Research Council (ERC) funding with the bibliometric output of successful grantees.

Design We analyzed publicly available data on the cohort of 2007-2009 ERC grantees in the Life Sciences domain (N = 355) for the Starting Grant (StG; n = 184) and the Advanced Grant (AdG; n = 171). The numbers of articles/reviews and citations in Scopus and Web of Science (WoS) were extracted for 5-year periods before and after the grant award. The Mann-Whitney U test was used for comparisons.

Results There were more male grantees (291 of 355 [82%]), both for the StG (144 of 184 [78%]) and the AdG (147 of 171 [86%]) ($P = .93$). The StG recipients published 2542 articles indexed in Scopus before the grant award and 4086 articles indexed in Scopus after the grant award, and they published 2476 articles indexed in WoS before the grant award and 3901 articles indexed in WoS after the grant award. The AdG recipients published 7448 articles indexed in Scopus before

the grant award and 8624 indexed in Scopus after the grant award, and they published 7197 articles indexed in WoS before the grant award and 8382 articles indexed in WoS after the grant award. The StG recipients had a significantly greater increase in the number of publications after the award (**Table 33**). Most of the publications by both grantee groups were in Scopus first-quartile journals (median percentage for StG recipients, 93% [95% CI, 92%-94%]; median percentage for AdG recipients, 92% [95% CI, 81%-94%]). The mean numbers of citations per publication were similar for both the StG and AdG recipients and did not change over time. The percentage of publications with the grantee as last author significantly increased for StG recipients and decreased for AdG recipients after the grant award. There were no gender differences for StG recipients; female AdG recipients had significantly fewer publications indexed in Scopus than did male AdG recipients after the grant award (median difference, -3.0 [95% CI, -6.5 to 3.3] vs 4.0 [95% CI, 2.0-7.0]; $P = .006$) but more last authorships indexed in Scopus (median difference, 3.6 [95% CI, -1.8 to 8.2] vs -0.1 [95% CI, -0.1 to 1.0]; $P = .006$). Female but not male StG recipients outperformed AdG recipients in publication output and last authorships (data not shown). There was no difference between StG and AdG recipients in the mean publication cost from the grant (total grant funding divided by number of publications after grant award: €63,000 [95% CI, €52,800-€70,300] vs €56,900 [95% CI, €50,000-€62,500]; $P = .08$).

Conclusions European Research Council funding to StG recipients was associated with increased numbers of publications and last authorships on these publications. An important limitation of our study was the lack of a control group of unsuccessful ERC grant applicants, but the data were not publicly available. It is not clear how to best measure the productivity of AdG recipients because their publication and citation output did not change with ERC funding.

Table 33. Publications (Articles and Reviews) by Successful ERC Starting and Advanced Grant Recipients and Citations to These Publications in WoS and Scopus 5 Years Before and After Grant Award

Database	Median (95% CI) ^a						P Value ^b
	Starting Grant Recipients (n = 184)			Advanced Grants Recipients (n = 171)			
	Before Grant Award	After Grant Award	Difference	Before Grant Award	After Grant Award	Difference	
No. of publications							
Scopus	11.0 (10.0 to 13.0)	20.0 (17.0 to 22.0)	7.0 (6.0 to 8.2)	33.0 (29.0 to 38.0)	37.0 (32.0 to 43.0)	3.0 (0.0 to 4.4)	<.001
WoS	11.0 (10.0 to 12.8)	19.0 (16.0 to 20.8)	6.0 (5.0 to 8.8)	31.0 (27.0 to 36.0)	37.0 (31.2 to 41.0)	2.0 (0.0 to 5.0)	.004
No. of citations per publication							
Scopus	19.3 (17.1 to 21.7)	15.6 (13.7 to 17.9)	−2.8 (−1.0 to −4.7)	20.8 (19.2 to 22.4)	18.7 (17.1 to 20.5)	−1.1 (−0.1 to −2.7)	.03
WoS	18.1 (16.4 to 20.9)	15.3 (13.0 to 17.7)	−3.1 (−4.8 to −1.7)	19.8 (19.1 to 21.7)	17.5 (16.6 to 19.8)	−1.6 (−3.5 to −0.4)	.07
% of Publications as last author							
Scopus	18.8 (13.5 to 25.0)	52.1 (48.7 to 56.6)	21.3 (16.4 to 21.3)	50.0 (46.9 to 52.9)	48.3 (43.8 to 51.1)	−4.1 (−6.3 to −0.4)	<.001
WoS	19.1 (15.4 to 25.0)	51.6 (47.5 to 54.9)	24.4 (16.8 to 30.0)	48.4 (46.2 to 52.0)	49.5 (43.6 to 51.0)	− 3.9 (−6.7 to −1.7)	<.001

Abbreviations: ERC, European Research Council; WoS, Web of Science

^aFor the purpose of this study, the grant award year (n) was considered the year of the call for proposals, as published in the respective ERC Work Programmes. The time span analyzed corresponded to the years n - 4 to n and n + 2 to n + 6 for the periods before and after the grant award, respectively. The ERC grant duration is about 5 years, depending on the project life cycle (in this set, 61% of grants lasted for 5 years, and 98% lasted for 4 to 6 years).

^bComparison of median differences, using the Mann-Whitney U test for independent samples.

¹Research Executive Agency, European Commission, Brussels, Belgium; ²Research Office, University of Split School of Medicine, Split, Croatia; ³Department of Research in Biomedicine and Health, University of Split School of Medicine, Split, Croatia, ana.marusic@mefst.hr

Conflict of Interest Disclosures: Dr Marušić is member of the Peer Review Congress Advisory Board but was not involved in the review or decision for this abstract. No other disclosures reported.

TUESDAY

Determining the Appropriateness of *Pediatrics* Case Reports Citations

Bryan A. Sisk,¹ Griffin S. Collins,² Claire Dillenbeck,³ J. Jeffrey Malatack⁴

Objective To determine the types of peer-reviewed articles (eg, original study or review) that cite *Pediatrics* case reports and the appropriate or inappropriate manner in which they are cited.

Design The 20 most highly cited *Pediatrics* case reports of the 381 published between January 2011 and April 2016 were identified. All English-language articles that referenced these case reports were obtained and analyzed for the appropriateness of the case report citation. We considered 2 types of appropriate citations: referring to the original publication specifically as a case report or citing the case report as background general knowledge. We considered an inappropriate citation as using the case report as proof of causal inference, proof of mechanism of pathogenesis or treatment, and when the citation was irrelevant to claims being supported. "Original studies" were defined as articles reporting original data, excluding case reports. Two authors independently coded all citations. Agreement was good for appropriateness vs inappropriateness ($\kappa=0.86$) and for further classification of appropriate citations ($\kappa=0.76$).

Results The top 20 case reports were cited by 479 articles (median, 24 citations per case report), accounting for 24.4% of all case report citations. For further analysis, articles were excluded if they were written in a non-English language (n=36), unable to be obtained (n=3), or erroneously included in our search (n=1). These remaining 440 articles were published in 281 unique journals. Most articles (83.4%, n=367) appropriately cited case reports, of which 53.4% (n=196) referred to the study specifically as a case report and 46.6% (n=171) used the case report to support general knowledge. For inappropriate citations, 63.3% (n=50/79) used case reports as proof of causal inference, 15.2% (n=12/79) used case reports as proof of mechanism of pathogenesis or treatment, and 21.5% (n=17/79) were irrelevant citations. Inappropriate citations of case reports were published in 60 unique journals, which ranged from national to international, from low to high reported Impact Factors. Case reports were most commonly cited by review articles (38.7%, n=170) and original studies (30.9%, n=136).

Conclusions Most of the more highly cited *Pediatrics* case reports reviewed were cited appropriately. These top 20 case reports were most commonly cited by review articles and original studies, suggesting that case reports support general knowledge and development of original studies. This study is limited by the narrow subset of case reports analyzed, which may limit generalizability. Further study should investigate whether case reports from other major journals are cited appropriately.

¹Department of Pediatrics, St. Louis Children's Hospital, St. Louis, MO, USA, sisk_b@kids.wustl.edu; ²Department of Pediatrics, St. Louis Children's Hospital, St. Louis, MO, USA; ³College of Arts and Sciences, Emory University, Atlanta, GA, USA; ⁴Department of Pediatrics, Thomas Jefferson University Sidney Kimmel Medical School, Wilmington, DE, USA

Conflict of Interest Disclosures: Bryan Sisk is the editorial fellow for *Pediatrics*. Jeffrey Malatack is the section editor for Case Reports in *Pediatrics*. No other potential conflicts of interest were reported.

Funding/Support: None reported.

Additional Contributions: We would like to acknowledge Lauren Yaeger for her support in developing search strategies for this study. Also, we would like to thank Dr Alex Kemper and Dr Lewis First for their continued support with revisions.

Conflict of Interest

MONDAY

Reporting of Conflicts of Interest of Panel Members Formulating Clinical Practice Guidelines in Anesthesiology: A Cross-sectional Study

Damien Wyssa,¹ Martin R. Tramèr,^{1,2} Nadia Elia^{1,2}

Objective Previous reports suggest that conflicts of interest (COIs) of panel members formulating clinical practice guidelines are common. Because COIs (both academic and financial) may bias the recommendations, solutions have

been proposed and implemented, such as the exclusion of panel members with important COIs or involvement of a methodologist without such COIs. However, little is known about the reporting of these procedures in the published guidelines. Our aim was to describe the way COIs are reported among anesthesia guidelines.

Design In this cross-sectional analysis of all guidelines published in 5 anesthesiology journals from 2007 to 2016, we examined the number and proportion of guidelines that (1) reported and described the potential influence of the guideline sponsor; (2) reported on individual panel members' COIs in the published guideline, and (3) did so in a clearly identified distinct paragraph; (4) included only panel members declaring a lack of COIs; (5) included a chair panelist declaring a lack of COIs; and (6) described procedures taken to minimize the risk of biases related to a panel member's COIs.

Results Our search strategy identified 76 publications, of which 66 met our inclusion criteria. Seven guidelines (11%) reported that they had received no funding, 8 (12%) reported having received funding without explanation of the potential impact of the sponsor on the recommendations, 2 (3%) reported funding with an explanation on the influence it could have had on the development of the guideline, and 49 (74%) made no statement regarding the funding of the guideline. Thirty-four guidelines (52%) reported on the panel member's COIs in the published report; 15 of 66 (23%) did so in a distinct paragraph with a subtitle mentioning the term "interest." Ten guidelines (11%) included only the panel member's statement declaring no COIs. Chairs of 35 guidelines (53%) were identified, of whom 8 reported a lack of COIs. Finally, 2 guidelines (0.3%) described measures taken to decrease the risk of biased recommendation related to the panel member's COIs.

Conclusion Although COIs of panel members are reported in about half of the published guidelines in anesthesia, the description of the COI and its potential influence on the guideline recommendation remains poorly documented. Standardized ways of reporting COIs of guideline panel members are required.

¹Division of Anesthesiology, Department of Anesthesiology, Pharmacology, and Intensive Care Medicine, Geneva University Hospitals, Geneva, Switzerland, martin.tramer@hcuge.ch; ²Faculty of Medicine, University of Geneva, Geneva, Switzerland; ³Institute of Global Health, University of Geneva, Geneva, Switzerland

Conflict of Interest Disclosures: Dr Tramèr is a Peer Review Congress Advisory Board Member but was not involved in the review or decision for this abstract.

TUESDAY

Physician Journal Editors and the Open Payments Program

Victoria S. S. Wong,^{1,2} Lauro Nathaniel Avalos,³ Michael L. Callahan³

Objective Open Payments is a federal program in the United States that requires reporting of medical industry payments to physicians. We examined this data to assess industry payments to physician journal editors, hypothesizing that physician journal editors would have a low rate of financial conflicts of interest (COI).

Design This was a retrospective study. We chose the top 5 representative and highly-cited clinical journals within each of these medical categories/specialties: general medicine, neurology, surgery, cardiology, and psychiatry. We systematically reviewed mastheads of 25 journals, identifying “top tier” editors who were considered senior within the editorial hierarchy. Our inclusion criteria aimed to target editors who were directly responsible for making manuscript decisions and avoid members who did not directly handle manuscripts. We were unable to confirm each editor’s role with their journals. We identified US-based physician editors and searched for industry payments to them using the publicly available Open Payments search tool. Data collected included general and research payments and ownership/investment records from August 1, 2013 (the start of the reporting requirement) to December 31, 2015.

Results Of 351 “top tier” editors of 25 journals, 246 (70%) met inclusion criteria as physician editors based at a US institution (mean, 9.8 editors per journal; SD, 9.8; range: 1-26). Of these, 160 (65%) received industry payments of any kind during the 29-month period. Eighty editors (33%) received direct payments (not to their institution) of \$5000 or more within a year, which met the threshold designated by the National Institutes of Health (NIH) as a significant financial interest (SFI). One hundred forty-nine editors (61%) received general industry payments (**Table 34**). The mean (SD) general industry payment to physician editors was \$37,225 (\$128,545) for the 29-month period while the median was \$2564. The mean (SD) total research payment made directly to physician editors was \$12,493 (\$34,710) with a median of \$1075. The mean (SD) research payment to the institution where physician editors were named as principle investigator on a research project (associated research payments) was \$105,283 (\$176,650) with a median of \$25,256. An additional \$12,766,532 paid over 3 ownership/investment transactions was reported; the bulk of this was in a single \$12,736,276 declaration of stock ownership, held by an immediate family member.

Conclusions Median direct industry payments to physician journal editors are generally low and do not surpass the SFI threshold designated by the NIH, suggesting overall low levels of financial COI. However, there are outliers, as evidenced by the high standard deviations from the mean values. Editor financial COI declarations may be appropriate given the extent of influence editors have on the medical literature.

*The Queen’s Medical Center Neuroscience Institute, Honolulu, HI, vwongmd@gmail.com; ²University of Hawaii, Honolulu; ³University of California, San Francisco, CA, USA

Table 34. Payments Received by United States Physician Editors from August 1, 2013 to December 31, 2015

Characteristic	General Payments	Total Research Payments	Associated Research Payments
No. (%)	149 (60.6)	36 (14.6)	70 (28.5)
Mean (SD)	\$37,225 (\$128,545)	\$12,493 (\$34,710)	\$105,283 (\$176,650)
Median	\$ 2564	\$1075	\$25,256

Conflict of Interest Disclosures: Dr Callaham receives fees from the American College of Emergency Physicians to edit their peer reviewed journal. No other disclosures were reported.

Funding/Support: None reported.

MONDAY

Effect of Different Financial Competing Interest Statements on Readers’ Perceptions of Clinical Educational Articles: A Randomized Controlled Trial

Sara Schroter,¹ Julia Pakpoor,² Julie Morris,³ Mabel Chew,⁴ Fiona Godlee⁵

Objective Declaration of conflicts of interest (COIs) is now the norm for research published in medical journals, with awareness that such interests may influence findings and reader perceptions. Less attention has been focused on the role of COIs in education articles intended to guide clinical practice. We investigated how different COI statements affect clinical readers’ perceptions of education articles.

Design We invited UK physicians in the British Medical Association membership database, by email, to participate in a research project in January 2016. One-third were general practitioners, one-third junior physicians, and one-third consultants. Volunteers were blinded and randomized to receive a shortened version of 1 of 2 clinical reviews (on gout or dyspepsia) previously published by *The BMJ* and considered to be of interest to a wide clinical audience. Each review was assigned 1 of 4 possible COI statements, and each review was identical except for permutations of the COI statement. After reading the review, participants completed an online questionnaire rating their confidence in the article’s conclusions (primary outcome), its importance, their level of interest in the article, and their likeliness to change practice after reading it. Blinded factorial analyses of variance and analyses of covariance were carried out to assess the influence of each review and type of COI on outcomes.

Results Of 10,889 physicians invited to participate, 1065 (9.8%) volunteered. Of these volunteers, 749 (70.3%) completed the survey. Analysis of covariance adjusting for age, sex, job type, and years since qualification showed no significant difference between the groups in participants’ confidence in the article (gout: $P = .32$, dyspepsia: $P = .78$) or their rating of its importance (gout: $P = .09$, dyspepsia: $P = .79$) (**Table 35**). For the gout review, participants rated articles with advisory board COI as significantly less interesting than those with no COI ($P = .02$ with Bonferroni

Table 35. ANCOVA Analysis of the Level of Confidence in the Articles by Intervention Group Adjusting For Age, Sex, Job Type, and Years Since Qualification

	COI Allocation Group, Mean (95% CI)				P Value
	Advisory Board ^a	Honoraria and Travel ^b	Research Funding ^c	None ^d	
Gout Article					
No.	90	99	93	90 ^e	
Primary outcome					
Level of confidence in conclusions drawn ^f	7.0 (6.7-7.4)	7.1 (6.8-7.5)	7.4 (7.1-7.8)	7.4 (7.0-7.8)	.32
Secondary outcomes					
Importance of article ^f	6.4 (6.1-6.8)	6.9 (6.6-7.3)	6.7 (6.4-7.1)	7.0 (6.6-7.4)	.09
Level of interest in article ^f	6.2 (5.9-6.6)	6.7 (6.5-7.0)	6.5 (6.2-6.9)	7.0 (6.7-7.4)	.03 ^g
Dyspepsia Article					
No.	93	100	95 ^e	87	
Primary outcome					
Level of confidence in conclusions drawn ^f	6.2 (5.8-6.6)	6.2 (5.8-6.6)	6.1 (5.7-6.5)	6.4 (6.0-6.8)	.78
Secondary outcomes					
Importance of article ^f	6.5 (6.2-6.9)	6.3 (6.0-6.7)	6.3 (5.9-6.7)	6.3 (5.9-6.7)	.79
Level of interest in article ^f	6.0 (5.6-6.4)	5.9 (5.5-6.3)	5.8 (5.4-6.2)	5.8 (5.4-6.2)	.83

Abbreviations: ANCOVA, analysis of covariance; COI, conflict of interest.

^aWe have read and understood the *BMJ* policy on declaration of interests and declare the following: DF is funded by a NIH clinician scientist award; SN receives no specific funding; JB has received fees from Jenka Pharmaceuticals for consultancies and being an advisory board member.

^bWe have read and understood the *BMJ* policy on declaration of interests and declare the following: DF is funded by a NIH clinician scientist award; SN receives no specific funding; JB has received honoraria and travel expenses from Jenka Pharmaceuticals for lecturing at a conference.

^cWe have read and understood the *BMJ* policy on declaration of interests and declare the following: DF is funded by a NIH clinician scientist award; SN receives no specific funding; JB has received research funding from Jenka Pharmaceuticals.

^dWe have read and understood the *BMJ* policy on declaration of interests and declare the following: DF is funded by a NIH clinician scientist award; SN receives no specific funding; JB has no competing interests.

^eOne respondent did not give ratings for confidence, importance, or interest level.

^fOutcomes measured on 10-point Likert scales with high scores indicating high levels of confidence, importance, and interest.

^gAllocation group “none” had a significantly higher level of interest compared with allocation group “advisory board” ($P=.02$ with Bonferroni correction).

correction). Among participants indicating that they treat the condition and that the article’s recommendations differed from their own practice, there was no significant difference in likelihood to change practice between groups (gout: $P = .59$, $n = 59$; dyspepsia: $P = .56$, $n = 80$).

Conclusions Physicians’ confidence in educational articles was not influenced by the COI statements. Our study was limited by a low response rate and by only using 2 articles that may not have been of interest to all participants. Further work is required to determine if physicians do not perceive these COIs as important in educational articles or if they do not pay attention to these statements.

¹*The BMJ*, London, UK, sschroter@bmj.com; ²Oxford University Clinical Academic Graduate School, Oxford, UK; ³University of Manchester, Manchester, UK; ⁴*The BMJ*, London, UK; ⁵*The BMJ*, London, UK

Conflict of Interest Disclosures: Sara Schroter, Fiona Godlee, and Mabel Chew are employed by *The BMJ*; Julia Pakpoor was a Clegg Scholar at *The BMJ*; and Julie Morris is a statistics editor for *The BMJ*. Fiona Godlee is a member of the Peer Review Congress Advisory Board but was not involved in the review of or decision for this abstract.

Funding/Support: We received no external funding for this study.

Additional Contributions: We thank the editors in *The BMJ* education team (Cath Brizzell, Tony Delamothe, Giselle Jones, Navjoyt Ladher, Emma Parish, Alison Tonks, Sophie Cook) and Theo Bloom for their help in designing the study; the authors of the original Clinical Reviews for giving us permission to use their work; Emma Parish for editing these articles for use in the study; and Keith Bates, Data Analyst at the British Medical Association, for providing data about association members and help with sampling.

TUESDAY

Competing Interest Disclosures Compared With Industry Payments Reporting Among Highly Cited Authors in Clinical Medicine

Daniel M. Cook,¹ Kyle Kaminski²

Objective Medical journals seek to minimize bias and enhance research integrity by requiring authors to disclose competing interests. The International Committee of Medical Journal Editors guidelines require disclosing related financial interactions from the 36 months prior to manuscript submission. Recently, the Affordable Care Act required the

health care industry (eg, drug and device manufacturers) to report payments to physicians. This created the Physician Open Payments Database maintained by the Centers for Medicare and Medicaid Services. We compared conflict of interest information from the Physician Open Payments Database with that disclosed by article authors.

Design Cross-sectional data were extracted from Open Payments and from medical journals and qualitatively coded. We started with the “Most Highly Cited Authors in Clinical Medicine for 2015” list from the Web of Science (n=375). For the authors with institutional affiliation in the United States (n=208), Open Payments data for 2014 were searched. For those authors with reported payments (n=121), we obtained 3 research articles published by them in 2015 or later. We examined the competing interest disclosure statements in the articles for concordance with the Open Payments report. Each article was coded as 1 of 4 categories: (1) full disclosure (total concordance with Open Payments); (2) partial disclosure (some financial ties found in Open Payments declared but not all, eg, Merck declared but Pfizer omitted); (3) declared different relationships (listed industry ties not in Open Payments); or (4) declared nothing to disclose.

Results A total of 363 articles were coded for 121 authors with 3 articles per author. One hundred sixty articles (44%) declared no competing author interests; 124 articles (34%) had partial (incomplete) author disclosure; 39 articles (11%) disclosed different financial ties than those found in Open Payments; and 40 articles (11%) were coded as having full concordance with the Open Payments. From among the 121 authors, 4 had all 3 sampled articles in full concordance with Open Payments. All 3 articles from 27 authors (22%) found in Open Payments claimed to have nothing to declare.

Conclusions Most of the highly cited authors sampled have not fully disclosed payments from industry. Our findings are consistent with those of other studies of scientist self-disclosures among conference presenters, clinical guidelines authors, and publications within specific specialties. Competing interest disclosures rely on trust and common understanding about the purpose. Authors may not perceive some industry payments as relevant to a particular article. The new source of payments data allows verification of submitted disclosures, and therefore improves assessment of the medical literature.

¹School of Community Health Sciences, University of Nevada, Reno, Reno, NV, USA, dmcook@unr.edu; ²Interdisciplinary PhD Program in Social Psychology, University of Nevada, Reno, NV, USA

Conflict of Interest Disclosures: None reported.

MONDAY

Collaboration Between Industry and Academics in Clinical Vaccine, Drug, and Device Trials: A Survey of Academic Investigators

Kristine Rasmussen,^{1,2} Lisa A. Bero,³ Rita Redberg,⁴ Peter C. Gøtzsche,¹ Andreas Lundh^{5,6}

Objective Collaboration between academic investigators and industry may constrain academic freedom. This study aims to determine the level of influence that academic investigators and funders have on industry-funded vaccine, drug, and device trials from design to publication, and to determine whether investigators report any problems with the collaboration.

Design We conducted a survey of lead academic authors. We included the most recent 200 trials meeting our inclusion criteria. These were phase 3 and 4 vaccine, drug, and device trials with sole industry sponsorship, at least 1 academic author, and published in 1 of the top 7 high-impact general medical journals (ie, *NEJM*, *Lancet*, *JAMA*, *BMJ*, *Annals of Internal Medicine*, *JAMA Internal Medicine*, and *PLoS Medicine*). We searched the journals on March 26, 2017. Two authors (A.L. and K.R.) independently extracted data on trial characteristics from the included publications, including the role of academic authors, funders, and contract research organizations. A pilot tested survey was emailed to lead academic authors. Questions were designed to ascertain who undertook each stage of the trial, whether disagreements arose with the funder at any decision point, and who ultimately had control over trial design, conduct, analysis, reporting, and publication. We analyzed the survey data descriptively and compared information reported in trial publications with survey responses.

Results The 200 included trials were published between July 2014 and March 2017. We included trials from all selected journals except *PLoS Medicine* where no industry-funded trials met our inclusion criteria. Of the 200 articles analyzed, 176 (88%) were coauthored by industry funders. We found that 170 (85%) trials reported funder involvement in design and 152 (76%) academic author involvement, yet the funder's and academic author's involvement in the design was not mentioned at all in 9 (5%) publications. Statistical analysis involved the funder in 147 (74%) trials and the academic author in 97 (49%), but their involvement in the statistical analysis was not mentioned in 9 (5%) publications. Trial reporting involved the funder in 168 (84%) trials and academic authors in 191 (96%), yet their role in the reporting was not mentioned at all in 6 (3%). Contract research organizations were involved in the trial reporting in 114 (57%) publications. Survey results are being analyzed and will be presented.

Conclusions In industry-funded trials published in high-impact journals, academic investigators are involved in the design and reporting of most clinical trials, but to a lesser degree in statistical analysis compared with industry funders.

¹The Nordic Cochrane Center, Rigshospitalet, Copenhagen, Denmark, kristinersmssn@gmail.com; ²Department of Primary Care & Public Health, Imperial College London, London, UK; ³Charles Perkins Center and Faculty of Pharmacy, University of Sydney, Australia; ⁴UCSF Division of Cardiology, University of California, San Francisco School of Medicine, San Francisco, CA USA; ⁵Center for Evidence-Based Medicine, Odense University Hospital and University of Southern Denmark, Odense, Denmark;

⁶Department of Infectious Diseases, Hvidovre Hospital, Hvidovre, Denmark

Conflict of Interest Disclosures: Lisa Bero and Peter C. Gøtzsche are members of the Peer Review Congress Advisory Board but were not involved in the review or decision for this abstract. No other conflicts are reported.

TUESDAY

Accuracy, Transparency, and Conflict of Interest in Medical Journal Drug Advertisements

James R. Scott,^{1,2,3} Mark Gibson,¹ Rebecca S. Benner³

Objective Medical journal editors have been leaders in criticizing financial conflicts of interest and supporting transparency in industry-sponsored research studies. Paradoxically, many medical journals continue to derive substantial income from their own advertisements for prescription medications. The purpose of this study was to evaluate the accuracy and medical value of these advertisements.

Design In this prospective observational study, we assessed all pharmaceutical advertisements for prescription drugs in the *New England Journal of Medicine*, *JAMA*, and *Annals of Internal Medicine* from May 2016 through October 2016. Two investigators independently reviewed the advertisements, and any disagreements were resolved by consensus. Outcomes examined included claims of efficacy; reporting of adverse effects; number, accessibility, and quality of references; and price of the drug.

Results Forty-two unique advertisements for 39 different drugs were identified among 190 total advertisements. Twenty-five of the 39 drugs (64%) were advertised more than once (range, 1-26). The retail price of 22 (56%) of these drugs was more than \$1000 for 1 month of treatment, and 7 (18%) cost more than \$10,000 per month. Most advertisements featured new drugs promoted through 1 to 3 pages of glossy and colorful attention-getting images followed by black-and-white package insert-like formats that contained detailed information about contraindications and adverse effects. Supporting references from the peer-reviewed medical literature were cited in 22 of the 42 different drug advertisements (52%). Forty-seven of 53 studies referenced (89%) were coauthored by individuals who had financial ties to the drug manufacturer, and they were also usually sponsored by the company. Twenty advertisements listed only prescribing information or data on file. Requests for data on file were successful in only 4 of 13 cases (31%).

Conclusions This study shows that contemporary pharmaceutical advertisements in major American medical journals promote expensive new drugs and do not provide sufficient information for review of claims made. Journal advertising that fails to foster dissemination of evidence-based knowledge and cost-effective patient care warrants reevaluation by publishers and editors.

¹Departments of Obstetrics and Gynecology, University of Utah School of Medicine, Salt Lake City, UT, USA, james.scott@hsc.utah.edu; ²University of Iowa Carver College of Medicine, Iowa City, IA, USA; ³*Obstetrics & Gynecology*, American College of Obstetricians and Gynecologists, Washington, DC, USA

Conflict of Interest Disclosures: None reported.

Funding/Support: None reported.

Data Sharing

MONDAY

Data Sharing Policies in Scholarly Publications: Interdisciplinary Comparisons

Michal Tal-Socher,¹ Adrian Zideman^{1,2}

Objective Data sharing is promoted in different avenues, a powerful one being the scholarly publication process. The research examines interdisciplinary differences in journal data-sharing policies and, to a smaller extent, data-sharing policies of major publishers.

Design The websites of a sample of academic journals and the websites of major journal publishers were examined for information on policies with regard to data sharing (as retrieved from October 2015 to December 2015). The journal sample was selected from 15 disciplines, drawn from 5 main academic discipline categories: biomedical sciences (including life sciences and medicine); physical sciences (including natural sciences without life sciences); social sciences (including economics, psychology, and political science); arts and humanities; and formal sciences (including cognitive science and statistics). Each disciplinary journal listing is ordered according to Scientific Journal Rankings score, and 10 journals were selected from each discipline; the total number of journals sampled is 150. Unlike the usual research strategy in the field, we opted for a broader sample distribution among different disciplines at the expense of depth. For each journal, the journal's website was searched to see whether there is a text instructing authors on the sharing of data supporting the research. Two definitions of data sharing were adopted for this purpose: *enabling data sharing*, defined as the sharing of academic article-related research data on an open digital platform, and *strong data sharing*, where at least some types of data must be deposited for open sharing as a condition for publication.

Results The central results, relating to journal policies in the 5 discipline categories, are shown in **Table 36**. The results also show the importance of major publishers in promoting data sharing. The 5 leading journal publishers account for 56% of the journals in our sample that enable sharing (and 46% of the strong policy journals), although they publish only one-third of the sample journals. The paper concludes with a presentation of the current preferences for different data-sharing solutions in different fields (ie, specialized repositories, general repositories, or publishers' hosting area).

Conclusions Assuming that journal and publisher policies are an important indicator of actual data sharing, the results

Table 36: Data sharing: interdisciplinary comparisons

Discipline group	Individual disciplines		Journals enabling data sharing	Journals with strong data sharing
	No.	Discipline	No. (%) ^a	No. (%) ^a
Biomedical Sciences	4	Genetics, Neuroscience, Oncology, Pharmacology (medical)	27 (67)	12 (30)
Physical Sciences	3	Chemistry, Geology, Ecology	13 (43)	3 (10)
Social Sciences	3	Economics, Social Psychology, Political Science & Int. Relations	14 (47)	3 (10)
Arts and Humanities	3	Archeology, Music, History	8 (27)	1 (3)
Formal Sciences	2	Computer science, Statistics & probability	7 (35)	1 (3)
Total	15	NA	69 (46)	20 (13)

^aOut of the journals sampled for each discipline group

consolidate the notion of the primacy of biomedical sciences in the implementation of data-sharing norms, the lagging implementation in the arts and humanities, and similar levels of norms adoption in the physical and social sciences. The results also point to the overlooked status of the formal sciences, which demonstrate low levels of data-sharing implementation. However, other tools for encouraging data sharing exist that may be stronger than publication policies in less journal-centric disciplines.

¹Consultant, Committee on Publication Ethics, Harlestone, UK, zidera@mail.biu.ac.il; ²Economics Department, Bar-Ilan University, Israel

Conflict of Interest Disclosures: None reported.

Funding/Support: The research was funded by the Committee on Publishing Ethics and conducted within the framework of its Research Subcommittee (Chair: Adrian Ziderman).

Disclaimer: The conclusions and views expressed are those of the authors and not necessarily of COPE.

TUESDAY

Endorsement of Data Sharing by Authors and High-Impact Medical Journals in China: A Survey of Authors and Assessment of Journal Online Instructions

Yuanyuan Ji,¹ Limin Chen,¹ Xiuyuan Hao,² Ningning Wang,¹ Yalin Bao¹

Objective To demonstrate that sharing data will increase confidence and trust in the conclusions drawn from clinical trials and avoid unwarranted repetition. This study was conducted to assess the endorsement of data sharing by authors and medical journals in China.

Design An electronic questionnaire was distributed via email between February 1, 2017 and February 13, 2017, to 438 corresponding authors of *Chinese Medical Journal* who published articles in 2016. The questionnaire contained 4

questions: (1) Have you ever published articles in international journals covered by the Science Citation Index (SCI)?; (2) Did you know about the requirements for sharing raw data for clinical trials before taking this survey?; (3) Do you endorse data sharing for clinical trials? If not, please state the reasons; and (4) Have you ever shared raw data through a data repository? The information obtained from the responses was extracted and analyzed by the χ^2 test to determine data sharing awareness and behavior between authors who published articles in international journals and those who published only in Chinese journals. We also investigated the websites of 111 high-impact journals hosted by the Chinese Medical Association to review their instructions for authors and editorial statements.

Results A total of 247 authors (56.4%) of 438 responded. Of these, 132 authors (53%) had published in international journals while 115 (47%) had only published in Chinese journals. Eighty-eight authors (35.6%) authors reported that they knew about data sharing, 215 (87%) stated that they endorsed data sharing, and 29 (11.7%) had shared raw data. Compared with authors who had published articles in international journals, those who published only in Chinese journals showed significantly lower data sharing awareness rates and behavior; however, no significant difference was found between the 2 groups regarding the endorsement rate (**Table 37**). The authors who did not endorse data sharing were mostly concerned about the misuse of their research data. Only 2 of 111 journals (1.8%) suggested that authors should share raw data and instructed ways to deposit data consistent with the policies of the International Committee of Medical Journal Editors.

Conclusions Most authors in China endorse data sharing; however, there is not a high rate of data sharing awareness and behavior. Authors who have published in international journals showed higher rates of data sharing awareness and behavior than those who have published only in Chinese journals. Medical journals in China should make efforts to help foster data sharing.

¹*Chinese Medical Journal*, Chinese Medical Association, 42 Dongsi Xidajie, Beijing 100710, China, jiyy@cma.org.cn; ²*Chronic Diseases and Translational Medicine*, Chinese Medical Association, 42 Dongsi Xidajie, Beijing 100710, China

Conflict of Interest Disclosures: None reported.

Funding/Support: None reported.

Table 37. Comparison of Data Sharing Awareness and Behavior Between Authors

Items	Published in International Journals (N=132)	Published Only in Chinese Journals (N=115)	P Value
Authors, No.	132	115	NA
Awareness rate, No. (%)	57 (43.2)	31 (27.0)	.008
Endorsement rate, No. (%)	110 (83.3)	105 (91.3)	.06
Sharing behavior, No. (%)	21 (15.9)	8 (7.0)	.03

Abbreviation: NA, not applicable.

Dissemination of Information

TUESDAY

NEJM Quick Take Videos: A Survey of Authors and Readers

Rebecca Berger,¹ Ramya Ramaswami,¹ Karen Buckley,¹ Roger Feinstein,¹ Kathy Stern,¹ Timothy Vining,¹ Stephen Morrissey¹, Edward W. Campion¹

Objective To assess authors' and readers' opinions about the *New England Journal of Medicine's* Quick Takes videos (QTs).

Design Since 2013, the *New England Journal of Medicine* has been creating short video summaries (QTs) for selected original research articles and making them freely available on NEJM.org and through social media. Producing QTs involves a collaborative in-house team, including editorial fellows for script writing, medical illustrators for graphics and production, and physician editors for review of accuracy, quality, and scope. Through December 2016, 96 videos were produced. From December 2016 to February 2017, an author survey was emailed to corresponding authors of articles for which QTs had been produced. From December 2016 to February 2017, a reader survey was advertised on NEJM.org, NEJM Resident360, Facebook, and Twitter.

Results Of 95 authors contacted, 48 replied (a 51% response rate). Thirty-six (75%) replied that they were very satisfied, and 8 (17%) that they were very dissatisfied with their role in helping to create QTs. Nine authors (19%) indicated they would have preferred more or earlier involvement in the process. Forty-seven authors (98%) somewhat or strongly agreed that the QT accurately summarized their article and presented it in an engaging way. Authors have used QTs for explaining research to family and friends (24 [50%]), promoting their research (23 [48%]), explaining research to colleagues (19 [40%]), presenting their research (15 [31%]), teaching (14 [29%]), and explaining the findings to patients (13 [27%]). A total of 411 readers responded to the reader survey, with response numbers differing by survey question. Of 332 respondents, 279 (84%) reported they had seen NEJM QTs. Most (198 of 237 [84%]) reported that QTs were valuable or very valuable to their education. When asked, "Do you believe that Quick Takes represent the abstracts of the future?" 210 of 254 (84%) responded "Yes." Among 236 respondents, 76% reported watching QTs to learn about new research without reading the article, 49% to decide whether to read the article (115 of 238 respondents), 34% to introduce a report they plan to read, 32% for entertainment, and 24% to teach. After watching a QT, 54% of the 236 responded that they read the associated article "sometimes" and 23% "about half the time."

Conclusions Quick Takes videos are making research more accessible to readers. Author and reader survey responses suggest that QTs are used for a variety of purposes. Many readers believe that the short video summary represents the abstract of the future.

¹*New England Journal of Medicine* and NEJM Group, Massachusetts Medical Society, Boston, MA, USA, reberger@partners.org

Conflict of Interest Disclosures: Dr Campion is a member of the Advisory Board of the Peer Review Congress but was not involved in the review of decision for this abstract.

Funding/Support: All authors receive salary support from the Massachusetts Medical Society.

MONDAY

Age of Clinical Trial Data at the Time of Publication: A Systematic Review of Clinical Trials Published in 2015

John Welsh,¹ Yuan Lu,^{1,2} Sanket S. Dhruva,^{3,4} Behnood Bickdeli,^{1,5} Nihar R. Desai,^{1,2} Liliya Benchetrit,² Chloe O. Zimmerman,² Lin Mu,² Joseph S. Ross,^{1,3,6,7} Harlan M. Krumholz^{1,2,3,7}

Objective To determine the age of clinical trial data at the time of trial publication.

Design Cross-sectional analysis of all clinical trials published in 2015 in *Annals of Internal Medicine*, *BMJ*, *JAMA*, *JAMA Internal Medicine*, *Lancet*, and *New England Journal of Medicine*. We determined the mid point of data collection to publication (age of data), in which the data collection period is defined as the start of enrollment to the end of follow-up. We also determined the days required for enrollment (enrollment time) and from the final study close-out visit to publication (dissemination time). We conducted multivariable linear regression models to identify factors associated with older data as well as longer enrollment and dissemination times.

Results Among 341 clinical trials published in the 6 journals in 2015, 206 were drug trials (60.4%), 21 were device trials (6.2%), and 114 were trials of other interventions (33.4%). The median age of clinical trial data was 1032 days (interquartile range [IQR], 714.5-1408.5); 10% of trials represented practice from 5 years or more at the time of publication. Median enrollment duration was 798 days (IQR, 431.25-1285.5) or 1.37 days (IQR, 0.5-3.8) required per person enrolled. A median 451 days (IQR, 225-674) elapsed from final study close-out to publication with 60% of trials requiring more than 1 year to publish and 18.5% of trials requiring more than 2 years. In multivariable analyses, a larger number of patients, a smaller number of trial centers and authors, and results not updated on ClinicalTrials.gov were statistically significantly associated with delays in publication because of the increased time required throughout 1 or more parts of a trial's duration (**Table 38**).

Conclusions By the time of publication, clinical trials in high-impact journals represent clinical practice with a median age of 1032 days, with 10% of trials representing practice from 5 or more years ago. There is substantial time required to publish a trial after the final study close-out.

Table 38. Clinical Trial Characteristics Associated With Time from Final Study Close-out Visit to Publication^a

Characteristics	No. Days to Publication (95% CI)	P Value
Study Type		
Drug	Reference	
Device	-57.3 (-226.5 to 112.1)	.51
Other	-21.4 (-124.3 to 81.5)	.68
Patients enrolled (per 1000), No.	4.2 (0.2-8.3)	.04
Trial centers, No.	-0.4 (-0.8 to -0.1)	.02
Authors, No.	-6.2 (-10.2 to -2.1)	<.001
Trial location		
United States only	Reference	
International only	43.8 (-73.4 to 161)	.46
United States and international	40.3 (-96.2 to 176.8)	.56
Manuscripts with at least 1 author primarily affiliated with private industry		
No	Reference	
Yes	90 (-48.2 to 228.2)	.20
Was the trial stopped early?		
No	Reference	
Yes	-87 (-248.7 to 74.6)	.29
Was the ClinicalTrials.gov page updated with results?		
Yes	Reference	
Not updated	137.4 (6.6-268.2)	.04
Registered on a site other than ClinicalTrials.gov	116.5 (-61.3 to 294.3)	.20
	46.8 (-474.4 to 567.9)	.86
Favorability of findings relative to the studied intervention		
Favorable	Reference	
Unfavorable	126.4 (-84.6 to 337.3)	.24
Inconclusive	40.4 (-54.2 to 135)	.40
Funding source		
Government	Reference	
Nonprofit	15 (-127.1 to 157)	.84
Private industry	-167.7 (-339.7 to 4.3)	.06
Government, nonprofit	-79.6 (-219.4 to 60.2)	.26
Government, private industry	-25 (-225.4 to 175.5)	.81
Nonprofit, private industry	-14.6 (-252.3 to 223.2)	.90
Government, nonprofit, private industry	-155.2 (-348.3 to 38)	.11
None	-290.6 (-986.1 to 405)	.41

^aEstimates derived from multivariable linear regression models.

¹Center for Outcomes Research and Evaluation, Yale New Haven Health, New Haven, CT, john.welsh@yale.edu; ²Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT; ³Robert Wood Johnson Foundation Clinical Scholars Program, Department of Internal Medicine, Yale School of Medicine, New Haven, CT; ⁴Veterans

Affairs Connecticut Healthcare System, West Haven, CT; ⁵Columbia University Medical Center/New York-Presbyterian Hospital, New York, NY; ⁶Section of General Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT; ⁷Department of Health Policy and Management, Yale School of Public Health, New Haven, CT, USA

Conflict of Interest Disclosures: Drs Desai, Ross, and Krumholz are recipients of a research agreement from Johnson & Johnson through Yale University to develop methods of clinical trial data sharing. Drs Ross and Krumholz receive research support from Medtronic through Yale University to develop methods of clinical trial data sharing and through a grant from the US Food and Drug Administration and Medtronic to develop methods for postmarket surveillance of medical devices. Dr Ross receives research grant support from the Blue Cross Blue Shield Association. Dr Krumholz is the founder of Hugo, a personal health information platform, chairs a cardiac scientific advisory board for UnitedHealth, is a member of the advisory board for Element Science, and is a participant/participant representative of the IBM Watson Health Life Sciences advisory board.

Funding/Support: This study received no external funding or support.

TUESDAY

Publication and Dissemination of Results in Clinical Trials in Neurology

Anirudh Sreekrishnan,¹ David Mampré,¹ Cora Ormseth,¹ Laura Miyares,¹ Audrey Leasure,¹ Lindsay Klickstein,¹ Joseph S. Ross,¹ Kevin N. Sheth¹

Objective ClinicalTrials.gov is an online database used to register and track clinical trials. Previous studies of this database have revealed low publication rates and selective reporting of results for registered trials. This study sought to classify the types of neurology studies within this registry and characterize both publication rates and time to publication.

Design We performed a search on July 19, 2016, using the ClinicalTrials.gov registry category “nervous system disease” to identify completed interventional studies conducted in the United States between October 1, 2007, and July 1, 2014. We then used both ClinicalTrials.gov and SCOPUS, the largest peer review journal database, to classify the topic subcategory and publication status as of December 2016. The primary outcomes were publication of results in a scientific journal and the time from study completion to publication. Publication status was identified through ClinicalTrials.gov entries as well as systematically searching SCOPUS with key terms for the primary investigator, condition, and intervention. Secondary outcome was the dissemination of results through ClinicalTrials.gov.

Results In all, 2072 studies were identified. Most funding came from industry (50.7%; n = 1051) or other nongovernmental sources (54.1%; n=1121), with only 12.9% of studies receiving National Institutes of Health funding (n = 267) and 7.1% receiving other US federal funding (n = 148). Of the 10 subcategories, the most prevalent was “pain medicine,” which accounted for 28.3% (n = 586) of all studies, followed by “behavioral/neuropsychiatry” (17.7%;

n = 367) and “neuromuscular medicine” (13.8%; n = 285). The overall publication rate was 47.8% (n = 990), with subcategory rates ranging from 38.8% (69 of 172) for “sleep medicine,” to 54.0% (154 of 285) for “neuromuscular medicine.” The median (SD) time to publication was 25 (15.2) months, with significant differences in time to publication among subcategories ($P = .003$). “Vascular neurology” had the fastest time to publication (median [SD], 18 [13.7] months) and was significantly faster than “pain medicine” (25 [15.6] months, $P = .029$), “behavioral/neuropsychiatry” (27 [16.5] months, $P = .002$), and “epilepsy” (28 [17.0] months, $P = .01$) on follow-up testing. Results were reported through ClinicalTrials.gov for only 40.6% (n = 841) of all trials and for exactly 50% (495 of 990) of all published studies.

Conclusions Fewer than half of neurology studies registered on ClinicalTrials.gov are published in a scientific journal, with rates varying among subcategories. Time to publication was consistent among subcategories, except “vascular neurology,” which had a significantly shorter time to publication. Further research can identify the source of variation such as journal niche, funding source, influential investigators, and research competition. Similar to other cross-sectional studies of this registry, further emphasis should be placed on result dissemination of registered trials including additional requirements and oversight.

¹Department of Neurology, Yale University School of Medicine, New Haven, CT, USA, david.mampre@yale.edu

Conflict of Interest Disclosures: Dr Sheth has received research funding from Remedy Pharmaceuticals Inc, Stryker Clinical Research, and C.R. Bard Inc. Dr Ross receives research support through Yale University from Johnson and Johnson to develop methods of clinical trial data sharing, Medtronic Inc, the Food and Drug Administration to develop methods for postmarket surveillance of medical devices and to establish a Center for Excellence in Regulatory Science and Innovation at Yale and Mayo Clinic, Blue Cross Blue Shield Association to better understand medical technology evaluation, Centers of Medicare and Medicaid Services to develop and maintain performance measures that are used for public reporting, and Laura and John Arnold Foundation to support the Collaboration on Research Integrity and Transparency at Yale. Dr Ross reports that he is a member of a scientific advisory board for FAIR Health Inc. Dr Ross receive support from the Centers of Medicare and Medicaid Services to develop and maintain performance measures that are used for public reporting. Dr Ross is supported by grant Ko8 AG032886 from the National Institute on Aging and by the American Federation for Aging Research through the Paul B. Beeson Career Development Award Program. No other disclosures were reported.

Funding/Support: No external funding was provided for this study. Contributions for staff time and resources came from the Yale University Department of Neurology.

MONDAY

Disclosure of Results of Clinical Trials Sponsored By Pharmaceutical Companies

Slavka Baronikova,¹ Jim Purvis,² Christopher Winchester,^{2,3} Eric Southam,² Julie Beeso,² Antonia Panayi¹

Objective To evaluate disclosure of clinical trials sponsored by pharmaceutical companies.

Design We used TrialsTracker to identify interventional phase 2-4 clinical trials that were registered on ClinicalTrials.gov; completed between 2006 and 2015; sponsored by the top 50 pharmaceutical companies (defined by 2014 global sales using EvaluatePharma); and that had results disclosed by April 2017, where disclosure is defined as registered on ClinicalTrials.gov or published in articles indexed on PubMed. We report the proportion of trials with disclosed results overall; by company membership in the European Federation of Pharmaceutical Industries and Associations (EFPIA) and Pharmaceutical Research and Manufacturers of America (PhRMA) and by industry vs nonindustry sponsorship.

Results Among the top 50 companies, 31 (62.0%) met inclusion criteria and were represented in TrialsTracker: 25 (80.6%) were EFPIA/PhRMA members and 6 (19.4%) nonmembers (generally medical device, generic drug, and non-EU/US companies). Among 6235 trials registered and completed by these companies between 2006 and 2015, results were disclosed for 4761 (76.4%), with the proportion rising from 42.9% in 2006 to approximately 80.0% from 2008 onwards (**Table 39**). The proportion of trials with results disclosed was similar for those sponsored by PhRMA/EFPIA members (1361 of 5697 [76.1%]) and nonmembers (113 of 538 [79.0%]). Of all clinical trials identified in TrialsTracker, results were disclosed for 74.0% of all pharmaceutical-industry sponsors and 45.7% of non-industry sponsors.

Conclusions The pharmaceutical industry has disclosed the results of three-quarters of trials completed between 2006 and 2015. Because TrialsTracker excludes sources other than ClinicalTrials.gov (eg, company websites), this figure may be an underestimate.

¹Shire, Switzerland GmbH, Zug, Switzerland, sbaronikova-c@shire.com; ²Oxford PharmaGenesis Ltd, Oxford, UK; ³Department of Medicine, Pharmacy and Health, Durham University, Durham, UK

Conflict of Interest Disclosures: S. Baronikova is a consultant to Shire Switzerland GmbH. A. Panayi is an employee of Shire Switzerland GmbH. J. Beeso, J. Purvis, E. Southam, and C. Winchester are employees of Oxford PharmaGenesis Ltd,

Table 39. Proportion of Pharmaceutical-Sponsored Trials For Which Results Were Disclosed by April 2017.

Category, No.	2006	2007	2008 ^a	2009 ^a	2010 ^a	2011 ^a	2012 ^a	2013 ^a	2014 ^a	2015 ^a
Companies	29	29	31	31	31	31	29	30	30	27
Trials	620	780	953	799	690	646	542	531	460	214
Proportion with results disclosed, %	42.9	54.4	81.0	86.1	84.6	87.2	89.3	82.1	84.1	73.4

^aMandatory reporting as required by FDA Amendments Act, Section 801, 2007.

Oxford, England. C. Winchester owns shares in Oxford PharmaGenesis Holdings Ltd and AstraZeneca Ltd, and is a Director of Oxford PharmaGenesis Ltd and Oxford PharmaGenesis Holdings Ltd.

Funding/Support: Funding was provided by Oxford PharmaGenesis and Shire, employees of which reviewed and approved the draft text.

TUESDAY

Frequency and Format of Clinical Trial Results Disseminated to Participants: A Survey of Trialists

Sara Schroter,¹ Amy Price,^{1,2} Mario Malički,³ Rosamund Snow,^{1,4} Tessa Richards,¹ Mike Clarke^{5,6,7}

Objective Dissemination of research findings is central to research integrity and supports the translation of clinical knowledge into practice. This survey investigates the frequency and format of research dissemination to trial participants and patient groups and explores how patients are involved in determining the content and method of dissemination.

Design First authors of clinical trials indexed in PubMed and published in English in 2014-2015 were emailed and invited to complete a SurveyMonkey questionnaire.

Results Surveys were sent to authors of 19,824 trials; 3227 responses were received (16%). Of the 3227 trials, 2690 had human participants and 1818 enrolled individual patients. Among the 1818, 906 authors (50%) had asked patients if they wanted to receive results, 305 (17%) had involved or planned to involve patients in developing dissemination materials, and 295 (16%) had involved or planned to involve patients in identifying appropriate dissemination methods. Four hundred ninety-eight (27%) reported that they had already disseminated results to participants and another 238 (13%) planned to do so, 600 (33%) did not plan to, 176 (10%) were unsure, and 256 (14%) responded with “other” or did not answer. Of the 498 authors who disseminated results to participants, 198 (40%) shared academic reports, 252 (51%) shared lay reports, and 164 (33%) provided individualized study results (Table 40). Among the 1818 trials, 577 authors (32%) shared or planned to share results with patients outside their trial by direct contact with charities/patient groups, 401 (22%) via informal patient communities, 845 (46%) via presentations at conferences with patient representation, 494 (27%) via mainstream media, and 708 (39%) by publishing lay summaries online. Relatively few authors of the 1818 trials reported that dissemination to participants and other patient groups was suggested to them by institutional bodies: by research funders (314 [17%] suggested dissemination to trial participants and 252 [14%] suggested dissemination to other patient groups), and by ethical review boards (333 [18%] suggested dissemination to trial participants and 148 [8%] suggested dissemination to other patient groups). Author-reported barriers to dissemination included not having access to patient contact information, insufficient funding and support, inaccessible

Table 40. Type of Information Offered to Participants in Trials for Which Authors Had Already Disseminated or Planned to Disseminate Results

Type of Information ^a	No. (%) of Trials	
	Already Disseminated to Participants (n=498)	Planned to Disseminate to Participants (n=238)
Documents or presentations written for an academic/clinical audience	198 (40)	105 (44)
Full study report	56 (11)	24 (10)
Journal publication(s)	139 (28)	67 (28)
Academic summary	74 (15)	47 (20)
Documents or presentations prepared specifically for lay readers	252 (51)	114 (48)
Participants sent a lay summary	170 (34)	82 (35)
Lay summary posted on a website	48 (10)	39 (16)
Trial participants invited to attend workshop or meeting	82 (16)	23 (10)
Individualized study results, such as outcomes, scores, or analyzed data	164 (33)	49 (21)

^aAuthors could indicate multiple options.

patient groups (eg, deceased, vulnerable, or mobile), time interval between the study and publication, and lack of patient-centered dissemination training.

Conclusion Fewer than half of respondents had disseminated or planned to disseminate their results to patients and only half of those in language tailored to patients. Motivation to disseminate appears to arise within research teams rather than from institutional bodies. Multiple factors need to be understood and overcome to facilitate wider, more effective dissemination of research to patients.

¹The BMJ, London, UK, sschroter@bmj.com; ²Department of Continuing Education, University of Oxford, Oxford, UK; ³School of Medicine, University of Split, Split, Croatia; ⁴University of Oxford Medical School, Oxford, UK; ⁵Northern Ireland Methodology Hub, Belfast, UK; ⁶Northern Ireland Clinical Trials Unit, Belfast, UK; ⁷Centre for Public Health, Queen’s University Belfast, Belfast, UK

Conflict of Interest Disclosures: Sara Schroter, Rosamund Snow, and Tessa Richards are or were employed by *The BMJ*, which has a patient partnership initiative. Amy Price, Rosamund Snow, and Tessa Richards are or were patients with long-term medical conditions and committed to increasing the dissemination of results to patients. Amy Price is a Research Fellow at *The BMJ*. Mario Malički has no conflicts of interest. Mike Clarke has been involved in many clinical trials and systematic reviews both as an organizer of the dissemination of his own research and using material disseminated by others. He seeks funding for these trials and reviews as well as for research into methodology, including dissemination.

Additional Information: Rosamund Snow is deceased.

Funding/Support: No external funding was received for this study. Mario Malički is partially funded by Croatian Science Foundation grant IP-2014-09-7672 Professionalism in Health Care.

Acknowledgments: We honor Rosamund Snow, *BMJ* Patient Editor, who passed away before this work could be published. We applaud her insightful perspective, faithful diligence, and humor as we worked together on this research. We thank the 3 members of the

BMJ Patient Panel for their comments on the survey content and the volunteers who pilot tested the survey, helped to improve the survey questions, and commented on the survey design.

Editorial and Peer Review Process

MONDAY

Editorial Rejections in *Obstetrics & Gynecology*

Randi Y. Zung,¹ Rebecca S. Benner,¹ Nancy C. Chescheir²

Objective To describe the effectiveness of *Obstetrics & Gynecology*'s editorial rejection process and to assess whether the journal is eliminating potentially publishable manuscripts from consideration.

Design Using retrospective data collected from administrative records maintained by the editorial office staff, manuscripts proposed for editorial rejection by the journal's editors from March 2013 through December 2015 were reviewed. Per journal procedures, manuscripts proposed for editorial rejection were reviewed by a designated member of the editorial board to verify agreement with the proposed rejection. If the editorial board member disagreed, the manuscript was automatically sent for peer review. All proposed editorial rejections were reviewed a second time during a weekly conference by the editors for final approval prior to author notification. If there was disagreement between the editors, the manuscript was sent for peer review.

Results *Obstetrics & Gynecology* received 5664 unsolicited manuscript submissions from March 2013 through December 2015; 1123 (19.8%) of these were case reports and 4541 (80.2%) were other unsolicited article types, such as Original Research. Of these submissions, 1116 (19.7%) were proposed for editorial rejection (**Table 41**). Overall, the editorial board member disagreed with the editor's proposed rejection for 43 manuscripts (3.9%), and these manuscripts were sent for peer review. Thirty-nine of the 43 manuscripts (90.7%) were ultimately rejected, requiring 70 peer reviews. Four of the 43 manuscripts (9.3%) were accepted following subsequent revision, requiring 15 peer reviews. Overall, 0.76% (43 of 5664) of submitted manuscripts were saved from outright editorial rejection, but only 0.07% (4 of 5664) were eventually published in the journal. The outright editorial rejection of 1116 manuscripts during this period saved *Obstetrics & Gynecology* from requesting 2825 peer reviews.

Conclusions Manuscripts proposed for editorial rejection and saved by the editorial board veto are generally rejected following peer review. The editorial rejection process seems to identify submitted manuscripts that would not likely be accepted for publication; however, generalizability is limited owing to the small number of manuscripts examined. Furthermore, the absence of a comparison group does not allow the editors to assess for certain that this process is not eliminating potentially publishable manuscripts. The journal acknowledges that the editorial board member may be more likely to agree with the editors' initial proposals, potentially making editorial rejections less likely to be saved by a veto.

Table 41. *Obstetrics & Gynecology* Editorial Rejections

Manuscripts ^a	Mar 2013 to Dec 2013	Jan 2014 to Dec 2014	Jan 2015 to Dec 2015
Unsolicited manuscript submissions	1670	1914	2080
Sent for review	1527	1512	1509
Proposed for editorial rejection	143	402	571
Rejected without review	136	385	552
Sent for peer review but rejected	7	16	16
Sent for peer review but accepted	0	1	3

^aData are given as numbers. Manuscripts included were unsolicited and would typically undergo peer review (Original Research, Case Reports, Systematic Reviews, Current Commentary, Executive Summaries, Consensus Statements, Guidelines, Clinical Practice and Quality, Procedures and Instruments, and Personal Perspectives). Those unsolicited types that are editor-reviewed (eg, Letters to the Editors) were not included. Finally, only manuscripts assigned to the 3 main editors were included.

¹*Obstetrics & Gynecology*, Washington, DC, USA, rzung@greenjournal.org; ²*Obstetrics & Gynecology*, University of North Carolina–Chapel Hill, NC, USA

Conflict of Interest Disclosures: None reported.

TUESDAY

Implementation of a Peer Reviewer Probation and Performance Monitoring Program at *The Journal of Bone & Joint Surgery*

Marc Swiontkowski,¹ Christina Nelson¹

Objective *The Journal of Bone & Joint Surgery (JBJS)* has been in publication for more than 125 years. Because of its relatively high Impact Factor and reputation, the *JBJS* reviewer roster is large and subject to steady growth. However, similar to other publications, *JBJS* is plagued with uneven reviewer performance. In recent years, several measures have been enacted to refine the reviewer database. Our goal was to evaluate the initial effect of these measures.

Design To encourage better reviewer performance, the editorial department contacted all individuals who met any 1 of a variety of poor-performance indicators in January 2016 (N = 612). These indicators included declining at least 3 consecutive review invitations, a less than 50% review-completion rate, not completing any reviews within the past 3 years, or low review grades in 2015. Each reviewer was informed of his or her removal from the reviewer roster, and those who requested to remain as a reviewer were flagged as “on probation” in the manuscript tracking system. Editors also had an opportunity to retain specific reviewers in the database, and these reviewers were also placed on probation. The performance of each probationary reviewer was then monitored closely over the following year with the goal of improving performance and consequently removing the probationary designation. A stagnant or poorer performance after 1 year would lead to the removal of that reviewer from the roster without notification.

Results The data for the 126 reviewers placed on probation are shown in **Table 42**. This group comprised approximately

Table 42. Performance Data for Reviewers on Probation^a

Data	Preprobation (01/01/15- 12/31/15) (n = 100) ^b	Postprobation (01/01/16- 12/31/16) (n = 107) ^c
Total No. of declined invitations	212	97
Total No. of uninvited reviewers	73	65
Mean review grade for submitted reviews (range, 80-100)	93	94
Time to acceptance of review invitation, mean, d	1.7	1.5
Time from invitation acceptance to review completion, mean, d	14.6	11.9

^aA total of 126 reviewers notified in January 2016.

^bDenotes the number of probationary reviewers who received at least 1 invitation to review a manuscript prior to probation status notification. Reviewers are typically invited to review a manuscript every 2 to 3 months.

^cDenotes the number of probationary reviewers who received at least 1 invitation to review a manuscript during the year after probation status was established.

9% of the overall reviewer database. There was a more than 50% reduction in the number of declined reviews, with an overall average of 2.12 declined invitations per reviewer before the probation period began and an average of 0.91 declined invitations per reviewer during the year following program implementation. There was also a reduction in the number of reviewers who were uninvited owing to no invitation response. The quality of the reviews was unchanged, and the mean number of days to acceptance of an invitation was not statistically different. There was also a decrease in the number of days to complete the review.

Conclusions A reviewer probation program appears to be successful in improving reviewer performance in an established surgical journal. In tandem with this program, we hope that the introduction of additional resources and incentives such as the Elite Reviewers Program and individual thank-you emails to reviewers who provide high-quality reviews will further enhance overall reviewer performance at *JBJS*.

¹The Journal of Bone & Joint Surgery, Needham, MA, USA, cnelson@jbjs.org

Conflict of Interest Disclosures: None reported.

TUESDAY

Evaluation of a Software Tool and Process to Assist Authors With Reporting Ethical Requirements of Studies

Tracy Ronan,¹ Alice Ellingham¹

Objective Authors are increasingly required to provide in-depth information when submitting articles, including ethics statements for Human, Animal, and Field Studies (such as the name of the institutional review board, the patient consent type, and the clinical trial registration details and dates). In our experience, authors, especially nonnative English-language speakers, can find formulating these statements a complex process. Designed as an initial proof of concept, we aimed to discover if helping authors with generating ethics statements before submitting articles could

speed up the submission process and provide a clearer statement.

Design Ethics statements were collected over a 3-day period from 13 biomedical journals. Each statement was checked against the journal’s criteria and marked as passed or failed. All failed statements were returned to the authors as per the individual journal’s processes. These were tracked to capture the additional information provided by the author to generate a statement via www.ethicsgen.com. These statements were then checked against the journal’s criteria to determine whether the author would have provided an acceptable statement at submission having used the author’s tool. All statements (excluding 47 incomplete statements) were passed through www.ethicsgen.com and then the order was randomized. The first 340 statements were offered to 7 preagreed editors in a survey for them to choose either their preferred statement or “no preference.”

Results Of 488 statements, 86 (17.6%) would not have passed their journal’s ethics criteria at the stage of checking and were returned to the author. Of these, 16 have since been resubmitted and the generated ethics statement would have been accepted in these cases, demonstrating that the original submission would have passed the ethics check had the author had access to the tool before submitting. The other 70 statements will continue to be resubmitted by authors, but the time frame for this study was too short to include them. Four editors completed the survey, with 291 of the offered 340 statements receiving at least 1 response (**Table 43**).

Conclusion While these results suggest that aiding authors to produce a statement may help some authors, this limited study does not provide enough detail to determine what level of time saving this would give an editorial office. A more in-depth study over a longer period would enable us to evaluate time savings or other efficiencies. Further work could identify which statements might benefit from this author tool and suggest how to provide additional support to authors.

¹Editorial Office Ltd, Overton, Hampshire, UK, tracy.ronan@editorialoffice.co.uk

Conflict of Interest Disclosures: Ms Ellingham is a codeveloper of Ethics Gen.

Funding/Support: None reported.

Table 43. Editor Responses to Author vs Generated Ethics Statements

	Percentage of Votes	Outright Preference
Author statement, No. (%)	183 (42%)	129 (44%)
Ethics Gen, No. (%)	76 (18%)	50 (17%)
No preference, No. (%)	172 (40%)	112 (38%)
Total votes, No.	431 ^a	291

^aTotal number of votes received. Some statements received responses from more than 1 editor.

Peer Review in Croatian Open Access Scholarly and Professional Journals: A Cross-Disciplinary Survey

Ivana Hebrang Grgić,¹ Jadranka Stojanovski^{2,3}

Objective The peer review process in Croatian open access journals was investigated to identify journal practices related to peer review and editors' opinions on open peer review.

Design An online questionnaire with 39 questions was sent by email to 441 journal editors from the Hrčak repository of Croatian open access journals. We collected the data on discipline, acceptance rate, peer review type, guidelines for peer reviewers, number of reviewers in the editorial systems, duration of peer review process, ethical issues, and editorial freedom and integrity. The data was collected in February 2017.

Results Ninety-eight editors responded (22%; 43 from science, technology, and medicine [STM] and 55 from social sciences and humanities [SSH]). According to the data from Hrčak repository, 217 of 441 journals use a peer review process. Only 1 response was from a journal that did not use peer review; thus, the editors of the 97 journals using peer review represent 45% of the Croatian open access journals that use peer review. Editors reported publishing 1130 retracted articles in their journals in the last 5 years (mean [SD], 12 [30] per journal), but the numbers provided cannot be verified because the Hrčak repository has only 11 retracted articles published by 9 journals in the same period. The mean journal acceptance rate is higher in SSH disciplines (45%) than in STM (36%) (**Table 44**). The mean number of peer reviewers in editorial systems is higher in the fields of STM (336) than in SSH (107). Double-blind review is used by 87% (48) of SSH journals and 49% of STM journals. Detailed guidelines for reviewers are provided by 40% (22) STM journals and 40% (17) SSH journals. Other journals provide short guidelines or structured forms for the reviewers. Less than half of the journals address ethical issues, such as plagiarism or confidentiality, in their guidelines for reviewers. None of the journals use public posting of open peer reviews. Eighty-six percent of the STM editors and 80% of the SSH editors believe they have editorial freedom and integrity, and 11% and 14%, respectively, believe they do not.

Conclusions These findings are limited by the low response rates and should be interpreted with caution. Peer review in Croatian open access journals represented by those editors who did respond lacks transparency and globally accepted standards. The concepts of open peer review are not well known to these editors. Croatian journals have a high level of editorial integrity, but there is a need to raise awareness of the importance of the transparent guidelines for the reviewers.

¹University of Zagreb, Faculty of Humanities and Social Sciences, Department of Information and Communication Sciences, Zagreb, Croatia, ihgrgic@ffzg.hr; ²University of Zadar, Department of

Table 44. Results of the Survey of Peer Review Process in Open Access Croatian Journals

Variable	Social Sciences & Humanities (n=55)	Science, Technology, Medicine (n=43)
Acceptance rate, mean (SE)		
No. of submitted manuscripts per year	31 (3)	103 (19)
No. of published papers per year	14 (1)	36 (5)
Mean acceptance rate, %	45	35
Timeliness of publishing, mean (SE)		
No. of d from submission to the peer reviewers' acceptance	17 (2)	11 (1)
No. of d from peer reviewers' acceptance to the last peer review submitted	46 (3)	59 (5)
No. of d from the last peer review submitted to the editor's decision on acceptance or rejection	16 (2)	13 (1)
No. of d from editor's decision to publication	68 (5)	56 (8)
No. of reviewers, mean (SE)		
Reviewers in the editorial system	108 (14)	337 (98)
Type of peer review, No. (%)		
Double-blind review	48 (87)	21 (49)
Single-blind review (identity of author known)	5 (9)	22 (51)
Not blind (identities of author and reviewers are known during the review process)	2 (6)	0
Open peer review (open public posting of reviews)	0	0
Guidelines for reviewers, No. (%)		
Detailed guidelines	22 (40)	17 (40)
Short guidelines	26 (47)	14 (33)
Forms	7 (13)	10 (23)
Nothing	0	2 (4)

Information Science, Zadar, Croatia; ³Ruder Bošković Institute, Bijenička, Zagreb, Croatia

Conflict of Interest Disclosures: None reported.

Funding/Support: None reported.

MONDAY

Feasibility of a Randomized Controlled Trial Comparing Results-Blind Peer Review vs Standard Peer Review for Reducing Publication Bias in an Open Peer Review Journal

Katherine S. Button,¹ Anna Clark,² Tim Shipley,² Liz Bal²

Objective To assess the feasibility of conducting a randomized controlled trial comparing the effectiveness of results-blind review with standard peer review on reducing publication bias in *BMC Psychology*, an open access, open peer review journal. Publication bias is defined as an excess of positive relative to null findings in the published literature, particularly among studies of low methodological quality.

Design Single-arm feasibility study to estimate key design parameters for a full-scale trial. Parameters include proportion of positive, unclear, and null findings in the accepted articles (primary outcome) and optimizing the criteria for their classification; methodological quality of the published articles (secondary outcome) and design of measures to assess quality; willingness of authors and reviewers to participate; editorial decisions at each stage including reversal of acceptance-in-principle decisions; and time needed to collect and analyze data. We also assess acceptability of the intervention to authors, editors, and reviewers and optimize technical workflow to minimize the editorial workload required for subsequent BMC journals to participate in the trial. Results-blind review where decisions to “accept in principle” and “reject” are based on review of a partial manuscript containing the rationale and methods sections alone, with results and discussion omitted. Full versions of manuscripts “accepted in principle” are then rereviewed by the same reviewers, where the decision to publish can only be reversed if the results and discussion deviate unjustifiably from the stated aims and methods.

Results Recruitment commenced on December 1, 2016. Of the 50 research articles submitted since, 14 (28 %) have entered into the results-blind feasibility study, 6 of which have received a first decision (2 accept in principle, 3 revise, and 1 reject and resubmit), and 2 final decisions (1 accept). The mean (SD) time to first decision of 46 (10) days is marginally lower than that of the journal average for standard submissions (82 [11]). The recruitment target for the feasibility phase is 20 research articles.

Conclusions To date, 28% of authors are willing to participate in the study, and we are working to further improve uptake. Based on these preliminary data, editorial decisions, peer-review times, and acceptance rates are no worse and may be slightly better than the journal’s averages, providing support for the feasibility of the intervention. Work optimizing criteria for classifying publication bias outcome measures is ongoing.

¹Department of Psychology, University of Bath, Bath, UK, k.s.button@bath.ac.uk; ²BioMed Central, London, UK

Conflict of Interest Disclosures: Drs Clark, Shipley, and Bal are employees of BioMed Central. Dr Button declares no conflicts of interest.

Acknowledgments: We thank the useful feedback on the implementation of this results-free peer review trial from the BMC Psychology Editorial Board and the Research Integrity group (<http://www.biomedcentral.com/about/who-we-are/research-integrity-group>), especially Maria Kowalczyk. We also thank members of the Editorial Office, who have helped with the implementation as part of the peer-review workflow for *BMC Psychology*, especially Ruth Baker and Sanam Sadarangani.

MONDAY

A Scoping Review of the Roles and Tasks of Peer Reviewers in the Biomedical Journal Editorial Process

Ketevan Glonti,^{1,2} Daniel Cauchi,³ Erik Cobo,⁴ Isabelle Boutron,^{2,5} David Moher,⁶ Darko Hren¹

Objective The purpose of this scoping review was to systematically determine what is known about the role and tasks of peer reviewers of biomedical journals.

Design We searched the following 8 electronic databases: the Cochrane Library, Cumulative Index to Nursing and Allied Health Literature; Educational Resources Information Centre; EMBASE; MEDLINE; PsycINFO; Scopus, Web of Science for literature that included competency-related statements pertaining to the role and tasks of peer reviewers of biomedical journals. The database search was supplemented by a review of grey literature at individual journal websites.

Results We screened 23,176 bibliographic records and identified 184 potentially relevant full-text publications, 174 of which were editorials, containing 53 unique statements related to the roles and tasks of peer reviewers of biomedical journal articles. We grouped these statements into 7 themes: (1) tasks related to reviewing the manuscript (eg, evaluate and improve manuscript quality; identify and alert to flaws in research design; aid authors in revising their manuscript for resubmission elsewhere); (2) tasks related to the editorial process (eg, assist editorial decision making regarding manuscript significance, pertinence to the journal discipline, and acceptance or rejection; advise editor on clinical credibility and usefulness for readers and/or practitioners; communicate ethical concerns to editors); (3) gate keeping role in maintaining journal reputation and credibility by ensuring that only good science is widely disseminated; (4) obligation for timely review (eg, need to respond in a timely manner); (5) obligation of confidentiality (eg, not sharing information with colleagues and others; destroying manuscript following review); (6) responsibility to exercise integrity while criticizing honestly and constructively, without bias; and (7) moral obligation to review as exercise of good citizenship (eg, civic duty of being a member of the scientific community).

Conclusions We found considerable variation in expectations and descriptions of tasks, roles, and responsibilities of peer reviewers involved in the editorial process of biomedical journals. These outcomes provide insight into the extent and nature of existing literature in this area, possibly leading to a future typology.

¹Department of Psychology, School of Humanities and Social Sciences, University of Split, Split, Croatia, kglonti@unist.hr; ²Université Paris Descartes, Sorbonne Paris Cité, Paris, France; ³Ministry for Health, Malta; ⁴Statistics and Operations Research Department, Barcelona-Tech, UPC, c/Jordi Girona, Barcelona, Spain; ⁵Centre d'épidémiologie Clinique, Hôpital Hôtel-Dieu, Notre-Dame, Paris, France; ⁶Center for Journalology, Clinical

Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Canada

Conflict of Interest Disclosures: Isabelle Boutron and David Moher are members of the Peer Review Congress Advisory Board but were not involved in the review or decision for this abstract.

Funding/Support: Ketevan Glonti, Erik Cobo, Isabelle Boutron and Darko Hren were supported by a grant by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 676207.

Role of the Funder/Sponsor: The funders had no role in the study design, data collection and analysis, or preparation of the abstract.

MONDAY

A Survey of Chinese Medical Researchers' Awareness and Use of Author-Suggested Reviewers

Limin Chen,¹ Xiuyuan Hao,² Yuanyuan Ji,¹ Yalin Bao¹

Objective The increase in medical manuscript submissions in China has burdened editors of general medical journals with finding specialist peer reviewers. Author-suggested reviewers might be a choice for editors. However, this practice might create problems that can compromise the integrity and impartiality of the peer review process. This survey investigated the perceptions of Chinese authors regarding the authors-suggested reviewers for their work.

Design We distributed a questionnaire by email in January 2017 to 583 corresponding authors who had published manuscripts in the *Chinese Medical Journal* in 2015 and 2016. The questionnaire contained 4 questions in Chinese: (1) Did you know that authors might suggest the peer reviewers for their work? (2) Have you ever suggested the peer reviewers for your work when you submitted your manuscript? (3) Did you know that cheating or fraud may be committed by authors when identifying author-suggested reviewers? (4) Have you ever manipulated the reviewer(s) you suggested when you were submitting a paper as an author? The survey also included open-ended questions or space for respondents to comment and ask questions.

Results We received 325 (55.7%) valid responses. A total of 317 respondents (97.5%) reported knowing that authors could suggest reviewers. However, only 122 (37.5%) confirmed that they had suggested the reviewers for the manuscript they submitted. Among the respondents, 73 (22.5%) admitted knowing that the use of author-suggested reviewers could be a means for cheating or committing fraud. Two respondents (0.6%) confessed to having faked the peer reviewer's email or suggesting their own names as reviewers, 5 (1.5%) refused to answer the fourth question. The participants also raised questions, such as the following: As suggesting reviewers is time-consuming for authors, could this step be skipped at the submission stage? Why do editors fail to identify the truthfulness and reliability of author-suggested reviewers? Why are instructions for suggested peer reviewers not provided by journals?

Conclusions Most of the Chinese authors in this survey were aware they may suggest reviewers, but only one-third had actually done so. The author-suggested reviewer process remains useful, and editors should respect authors' rights to suggest peer reviewers. However, the process of suggesting reviewers needs to be fine-tuned. Detailed instructions for suggesting peer reviewers should be stated on journals' websites.

¹*Chinese Medical Journal*, Chinese Medical Association, Beijing, China, chenlimin@cma.org.cn; ²*Chronic Disease and Translational Medicine*, Chinese Medical Association, Beijing, China

Conflict of Interest Disclosures: None reported.

Funding/Grant Peer Review

MONDAY

Influences of Independent Peer Reviewer Scores on UK National Institute for Health Research Grant Funding Recommendations

Nicola McArdle,¹ Helen Payne,¹ Sheila Turner,¹ Jeremy C. Wyatt²

Objective Research funding boards expend much effort obtaining and considering peer reviews before making funding recommendations. While a reviewer's textual comments provide most value, their review score should be a good proxy for opinion regarding funding outcome. The United Kingdom's National Institute for Health Research (NIHR) is keen to improve review processes so we studied relationships between reviewer scores, funding board scores, and funding recommendations across 4 programs managed by a single NIHR coordinating center.

Design This is a cross-sectional study of reviewer and board scores for all full applications assessed in 2015. Both scores can range from 1 to 6 (1 indicates unfundable; 2 and 3, major changes necessary; 4 and 5, fundable with changes; and 6, fundable). Relationships between individual and mean reviewer scores and board scores per application were compared, and the influence of scores from different reviewer types (health economist, methodologist, clinician, patient, subject expert; number of funded applications to NIHR programs in this study) on board funding recommendations assessed using logistic regression and receiver-operator characteristic (ROC) curves.

Results The analysis is based on 1599 reviewer scores from 295 funding applications (54% successful). The number of reviews per application varied from 2 to 9 (mean [SD], 5.5 [1.2] reviews) and was not correlated with the amount of funding requested. Reviewers tended to be more generous than boards to applications scoring in the low- to mid-range. Logistic regression (based on decision to fund or not [265 applications]) showed a large increase in funding probability for each unit increase in mean reviewer score (relative risk, 7.0; 95% CI, 3.7-13.5) but no correlation with the number of reviewers per application; ROC curves using mean reviewer scores showed fair discrimination (area under ROC [AUROC],

0.77; 95% CI, 0.71-0.82) (**Figure 9**), which changed little from 4 to 7 or more reviewers per application. There was a trend towards increasing AUROC for individual reviewers receiving more NIHR grants. There were minor differences in AUROC by type of reviewer (0.60 for health economist or methodologist; 0.61, clinician; 0.64, patient representative; 0.66, subject expert). All appeared lower than the AUROC for mean reviewer score per application.

Conclusions Our analysis covering 4 NIHR programs shows large variation in the number of peer reviews informing recommendations. Our results suggest that sometimes 4 reviewers might suffice. The type of reviewer does not make much difference, but using the mean score from different types of reviewer instead of individual scores does.

¹National Institute for Health Research, University of Southampton, Southampton, UK, nicola.mcardle@nihr.ac.uk; ²Wessex Institute, University of Southampton, Southampton, UK

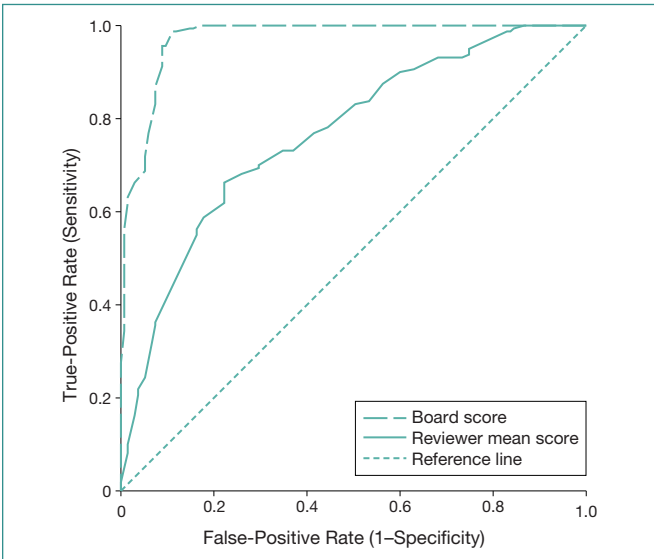
Conflict of Interest Disclosures: None reported.

Funding/Support: This research was supported by the National Institute for Health Research Evaluation, Trials, and Studies Coordinating Centre (NETSCC) through its Research on Research program.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the abstract.

Disclaimer: The views and opinions expressed are those of the authors and do not necessarily reflect those of the Department of Health or of the NETSCC.

Figure 9. Different Possible Cut Points for Fund or Reject Recommendation.



The area under curve for the Board score is 0.97 (95% CI, 0.95-0.99); Reviewer mean score, 0.77 (95% CI, 0.71-0.82).

TUESDAY

Influence of Evaluation Criteria on Overall Assessment in Peer Review of Project Grants Submitted to the Swiss National Science Foundation

Stéphanie Würth,¹ Katrin Milzow,¹ Matthias Egger¹

Objective The Swiss National Science Foundation (SNSF) supports basic science in all disciplines. External peer reviewers assess 3 evaluation criteria: (1) the scientific track record and expertise of applicants (track record); (2) the project’s scientific relevance, originality, and topicality (scientific relevance); and (3) the suitability of the methods and feasibility (methods and feasibility). The peer reviewers also provide an overall assessment. We aimed to identify the relative influence of the 3 evaluation criteria on the overall assessment.

Design We analyzed reviewer ratings for 7139 project proposals submitted from 2012 to 2015. A total of 21,902 reviews were obtained from experts from more than 80 countries. The external reviewers rated the 3 evaluation criteria and the overall assessment in 6 categories: outstanding, excellent, very good, good, average, and poor. We assigned scores from 1 (poor) to 6 (outstanding) to the ratings. We performed a multivariate linear regression analysis, with the score from the overall assessment as the dependent variable and the scores from the 3 criteria as the independent variables.

Results The mean score for the overall assessment was 4.46. Among the 3 criteria, the track record of the applicants received the highest average score (4.82), followed by scientific relevance (4.58) and the suitability of methods and feasibility (4.27). The linear multivariate regression analysis showed that the overall assessment was influenced most by the suitability of the methods and feasibility (b, 0.436; 95% CI, 0.430-0.442), followed by scientific relevance (b, 0.401; 95% CI, 0.394-0.408) and the applicants’ track record (b, 0.221; 95% CI, 0.214-0.228) (**Table 45**). There were interesting differences between fields of research (humanities and social sciences, mathematics, natural and engineering sciences, and biology and medicine); the suitability of the methods and feasibility was more important in biology and medicine, whereas the track record received more weight in mathematics and natural and engineering sciences.

Table 45. Results from Multivariate Linear Regression Analysis

Criteria	Coefficient	SE	95% CI
Constant	-0.298	0.013	-0.323 to -0.273
Scientific track record and expertise of applicants	0.221	0.004	0.214-0.228
Scientific relevance, originality, and topicality	0.401	0.004	0.394-0.408
Suitability of methods and feasibility	0.436	0.003	0.430-0.442

Abbreviation: SE, standard error.

Conclusions The greater weight given to the methods and feasibility and scientific relevance align with the principles of the project rather than career funding. Differences between fields of research exist and may reflect differences in research cultures.

¹Swiss National Science Foundation, Bern, Switzerland, stephanie.wuerth@snf.ch

Conflict of Interest Disclosures: Prof Egger is president of the Swiss National Science Foundation's National Research Council.

TUESDAY

Patient-Centered Outcomes Research Institute (PCORI) Methodology Standards to Improve the Design and Reporting of Research

Evan Mayo-Wilson,¹ Kelly Vander Ley,² Kay Dickersin,¹ Mark Helfand²

Objective The Patient-Centered Outcomes Research Institute (PCORI) began receiving funding applications in September 2011 and published Methodology Standards in November 2013 addressing issues related to research design and transparent reporting. PCORI requires that investigators of funded studies submit a draft final research report (DFRR) that is peer reviewed by an external team; after revision, in response to peer-review, the final report is published on the PCORI website. We sought to determine whether research described in DFRRs adheres to the PCORI standards.

Design Research funded before the standards were published (November 2013) was eligible for the study, as were DFRRs received by February 1, 2017. Thus, the standards were not in place when eligible studies were funded but were in place when DFRRs were submitted. We excluded PCORI-funded research for "improving methods for conducting PCOR." In 2017, we used the DFRRs to develop a 57-item checklist to assess adherence to the standards. Two authors (E.M.W. and K.V.L.) independently rated each eligible DFRR, compared disagreements, and resolved differences through discussion.

Results Among 31 eligible DFRRs, none adhered to all standards. Nonadherence was attributed to both incomplete reporting and to nonadherence with recommendations for study design. Examples of nonadherence in 5 areas of interest include (1) formulating research questions: most reports neither included nor cited a systematic review, and most did not include or cite a study protocol; (2) patient-centeredness: all reports mentioned patient engagement but none described engaging patients in all parts of their research as specified in the standards; (3) data integrity and rigorous analyses: most reports included patient-reported outcomes but few described how patient-reported outcomes would be interpreted by patients; (4) preventing and handling missing data: many reports did not use appropriate methods for handling missing data; and (5) heterogeneity of treatment effects: most reports examined heterogeneity, commonly using subgroup analyses,

but few studies were designed to conduct confirmatory tests for heterogeneity.

Conclusions Our results identify common limitations in research funded by PCORI prior to publishing the 2013 Methodology Standards. Further studies are needed to determine whether peer review of DFRRs improves adherence to the standards in the final research reports and whether peer reviewers consider nonadherence to be related to important limitations in interpretation. Observational studies might explore whether research funded after the standards were published is better designed and better reported.

¹Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA, evan.mayo-wilson@jhu.edu;

²Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, OR, USA

Conflict of Interest Disclosures: Kay Dickersin is a member of the Peer Review Congress Advisory Board but was not involved in the review or decision for this abstract. None of the authors was directly involved in any of the 31 eligible and included studies. During peer review of PCORI-funded research, the authors recused themselves from discussions about research conducted at their institutions (ie, Johns Hopkins University for Kay Dickersin and Evan Mayo-Wilson; Oregon Health & Science University for Mark Helfand and Kelly Vander Ley); these reports were included in the sample assessed by Evan Mayo-Wilson and Kelly Vander Ley for this study.

Funding/Support: This study was conducted under a contract to provide peer review services for PCORI research reports. Mark Helfand is the principal investigator of the Peer Review of PCORI's Funded Research, and the other authors are associate editors. All authors receive support from PCORI for reviewing research reports and for conducting research about the peer review process.

Role of the Funder/Sponsor: PCORI was not involved in conducting this study, but the study used research reports submitted as part of the PCORI peer review process. PCORI reviewed and approved this abstract for submission.

Acknowledgments: We thank Holly Somers and Kira Lesly, Oregon Health & Science University, for their assistance in identifying and organizing eligible research reports.

Integrity and Misconduct

TUESDAY

The Journal Project of the Russian Dissernet

Andrey Rostovtsev,¹ Alexei Kassian,² Vasilij Vlassov,³ Anna Abalkina,⁴ Larisa Melikhova⁵

Objective Dissernet is a voluntary organization of Russian scientists devoted to the identification of research misconduct, primarily plagiarism in dissertations. Over 4 years, and counting only multipage plagiarism, Dissernet identified and made public more than 6000 falsified dissertations. To further study the issue of misconduct, Dissernet inaugurated the Journal Project in 2016, with the aim of systematically identifying publication or research misconduct (plagiarism, duplicate publications, gifted and

stolen authorship, fake peer reviews, and other violations) in Russian scientific journals.

Design Russian journals were chosen from the Register of Scientific Journals of the Higher Attestation Commission of the Russian Ministry of Education and Science, where it is recommended that PhD candidates publish their papers as a means of working with officially recognized journals. Full text of all articles published in the selected journals was studied. All text mining was performed semiautomatically using special software, with review by study investigators for quality assurance.

Results As of February 2017, more than 3000 of 100,000 journal articles contained evidence of research misconduct. A subsample of 1000 articles with more than 80% text duplication contained evidence of plagiarism (38%), duplicate publications (37%), and fictitious authorship, including names of authors who do not exist (25%). Research misconduct was most frequently identified in publications relating to economics, law, and the pedagogical sciences. Misbehaving journals may be categorized as (1) predatory journals (high yield of plagiarism, distinguished by a large volume of short reports in different subject areas; sharp increase in number of published reports; absence of recognized expert scientists on the editorial board; or maximum cost saving on the editorial and publishing process, visible from poor editing, layout, and web presentation); (2) traditional journals (review process is formal or does not exist at all; editorial process is of low quality; low attractiveness to degree candidates because of low citation rate and other metrics); and (3) journals informally collaborating with the dissertation councils (“disser-collaborating” journals; these dissertation councils approve a high number of falsified dissertations containing plagiarism, and members of these councils are members of the editorial boards of these journals).

Conclusions Scientific periodicals in Russia are in a catastrophic state. We believe this is an outcome of the weakness of scientific expertise in society in general since the Soviet era. Moreover, the official formal criteria for publications, which are difficult to meet for the average degree candidate, may in part account for the pressure to manipulate publications.

¹Institute for Information Transmission Problems, Moscow, Russia;

²Institute of Linguistics of the Russian Academy of Sciences, Moscow, Russia; ³National Research University Higher School of Economics, Moscow, Russia, vlassov@cochrane.ru; ⁴Financial University under the Government of the Russian Federation, Moscow, Russia; ⁵T-Systems, Saint Petersburg, Russia

Conflict of Interest Disclosures: None reported.

Funding/Support: None reported.

MONDAY

Misuse of Received Manuscripts by Peer Reviewers: A Cross-sectional Survey

Darren Taichman,¹ Jill Jackson,¹ Deborah Cotton,¹ Cynthia Mulrow,¹ Jaya Rao,¹ Mary Beth Schaeffer,¹ Catharine Stack,¹ Sankey Williams,¹ Christine Laine¹

Objective How often peer reviewers use information from unpublished manuscripts in a manner inconsistent with the goals of peer review is not known. *Annals of Internal Medicine* recently experienced an egregious occurrence involving the plagiarism of an entire research study by a peer reviewer who had evaluated the manuscript for *Annals of Internal Medicine*. The reviewer subsequently published the study in another journal as his own work. We therefore aimed to assess peer reviewers' views and practices including their self-reported use of any information in manuscripts they reviewed.

Design All recipients of *Annals* manuscripts sent for external review in 2015 and 2016 were invited to complete an anonymous online survey between December 8, 2016, and January 17, 2017. Two reminder emails were sent.

Results A total of 1431 of 3275 invited reviewers (44%) returned the survey; 1068 of 1398 respondents (76%) reported working in an academic setting and 1249 of 1388 (90%) reported being involved in research. Nearly half indicated having reviewed and published more than 50 manuscripts and having mentored others in peer review. Reasons reported for agreeing to review included keeping up to date in a research field (957/1417 [68%]), a sense of obligation to peer review (1316/1417 [93%]), to plan one's own work (425/1417 [30%]), and to know what competitors are doing (190/1417 [13%]). One hundred sixty-nine of 1417 (12%) had agreed to review manuscripts from authors with whom they had conflicts of interest; of these, 61 (36%) did so without informing the journal's editor. One hundred fifty-three of 1413 (11%) showed manuscripts to colleagues without seeking permission. Twenty-six of 1414 (2%; 95% CI, 1%-3%) indicated having used the information in a reviewed manuscript for personal or academic benefit prior to the paper's publication. Such reported use included using what was learned to alter one's own research plans, speeding up journal submission of one's own work related to the subject of the manuscript being reviewed, and copying some part of the reviewed manuscript for one's own work.

Conclusions Trust that reviewers will treat manuscripts received for peer review as confidential communications is an essential tenet of peer review. Although self reported and of uncertain generalizability, these results suggest that breaches of this trust do occur. Larger studies involving multiple journals should be considered to assess the generalizability of these results and to inform targeted educational initiatives aimed at promoting the highest ethical standards among peer reviewers.

Conflict of Interest Disclosures: Christine Laine is a member of the Peer Review Congress Advisory Board but was not involved in the review or decision for this abstract. Jaya Rao reports that she has stock holdings/options in Eli Lilly and Pfizer. Catharine Stack reports that she has stock holdings in Pfizer and Johnson & Johnson.

MONDAY

A Survey of Knowledge and Perception of Plagiarism Among Chinese Authors and Reviewers

Pei-Fang Wei,¹ Xiu-Yuan Hao,¹ Yang Pan,¹ Wei-Zhu Liu,¹ Jing-Ling Bao,¹ Jun-Min Wei,¹ Yong-Mao Jiang¹

Objective For the past decade, plagiarism has become a growing problem among Chinese researchers; it has soiled the academic world, especially the medical community in China. In this survey, we sought to investigate knowledge and perceptions of plagiarism among Chinese authors as well as authors who were also reviewers and to clarify the standards for plagiarism.

Design In January 2017, a 10-item questionnaire was distributed by email to 252 authors of articles published in the *Chinese Medical Journal* (volume 128, issues 13-24) and 63 authors of articles published in *Chronic Diseases and Translational Medicine* (volumes 1 and 2). Among the 315 authors, 62 (19.7%) also served as reviewers. The questionnaire included questions about different forms of plagiarism (yes/no questions) and standards of plagiarism (multiple choice questions). Differences between the authors and reviewers were analyzed with the χ^2 test.

Results We received responses from 70.5% of authors (n = 222, including 47 who were also reviewers). Among the responders, 60.8% (135) regarded the overall similarity index of 20% to 30% (threshold) as acceptable; 55.4% (123) thought the duplication of a maximum of 20 to 30 words from someone else's work without references is plagiarism, and 31.1% (69) believed that the continuous duplication of a maximum of 100 words from other papers is plagiarism even if the references and quotation marks are included. Responses to the 10 questions about types of plagiarism are listed in **Table 46**. The χ^2 test indicated that there was no difference between authors and authors who were also reviewers in their knowledge and perceptions of plagiarism, except in their opinions about the similarity index ($P = .048$) and the provision of incorrect references ($P = .049$).

Conclusions There is a lack of agreement about the threshold for plagiarism among these Chinese authors, which implies that there is a need to establish a unified standard for researchers to follow. Chinese authors, including those also serving as reviewers, lack sufficient knowledge about plagiarism. Hence, special education programs should be provided to prevent plagiarism.

Table 46. Comparison of Answers Between Authors and Authors Who Are Also Reviewers

Question	Authors Who Are Also Reviewers (n = 47)	Authors (n = 175)	P Value
Acceptable overall similarity index (threshold)			
10%	8 (17.0)	17 (9.7)	.048
15%	12 (25.5)	29 (16.6)	
20%	6 (12.8)	46 (26.3)	
25%	10 (21.3)	20 (11.4)	
30%	9 (19.1)	44 (25.1)	
50%	2 (4.3)	19 (10.9)	
How many words of duplication from someone else's work without references is plagiarism?			
6	4 (8.5)	11 (6.3)	.82
15	8 (17.0)	44 (25.1)	
20	14 (29.8)	48 (27.4)	
30	14 (29.8)	47 (26.9)	
50	7 (14.9)	25 (14.3)	
Completely copying someone else's work with references but not quotation marks is plagiarism			
Yes	27 (57.4)	94 (53.7)	.65
No	20 (42.6)	81 (46.3)	
How many words of continuous duplication from other papers is plagiarism?			
50	8 (17.0)	35 (20.0)	.60
100	14 (29.8)	55 (31.4)	
150	8 (17.0)	18 (10.3)	
200	12 (25.5)	38 (21.7)	
300	5 (10.6)	29 (16.6)	
Restatement, paraphrase, and translation of someone else's work without citations is plagiarism			
Yes	45 (95.7)	152 (86.9)	.09
No	2 (4.3)	23 (13.1)	
Providing incorrect references is plagiarism			
Yes	15 (31.9)	84 (48.0)	.049
No	32 (68.1)	91 (52.0)	
Not inserting the references due to negligence is plagiarism			
Yes	35 (74.5)	123 (70.3)	.57
No	12 (25.5)	52 (29.7)	
Using the tables and figures from someone else's work without authorization, even though sources are indicated, is plagiarism			
Yes	20 (42.6)	85 (48.6)	.46
No	27 (57.4)	90 (51.4)	
Citing someone else's work in the core concept, key evidence, and important data of your paper is plagiarism			
Yes	13 (27.7)	76 (43.4)	.05
No	34 (72.3)	99 (56.6)	
Reusing the contents from your previous paper without references is plagiarism			
Yes	13 (27.7)	53 (30.3)	.73
No	34 (72.3)	122 (69.7)	

Conflict of Interest Disclosures: None reported.

TUESDAY

Post-retraction Citations in Korean Medical Journals

Sun Huh,¹ Hyun Jung Yi,² Hye-Min Cho,³ Soo Young Kim⁴

Objective The aim of this study was to identify post-retraction citations of articles in Korean medical journals indexed in the KoreaMed database and to investigate whether post-retraction citations depended on whether the retraction notice was made via PDF or the homepage. Although some studies have been conducted on the extent of post-retraction citations among PubMed-indexed articles, to our knowledge, no studies have evaluated whether this tendency differs among countries.

Design Retracted articles from the KoreaMed database were collected on January 28, 2016, and PDF files were obtained, along with citation information on the journal homepage. At the time of the study, KoreaMed contained 254,000 citations from 242 medical journals published in Korea. The linkage between the retracted article and the notice was the same as has been used for PubMed. The Web of Science Core Collection and Scopus were searched for post-retraction citations, which were defined as citations 1 year after the retraction, excluding retraction-related citations, considering the time to publication after submission. For each article, it was recorded whether the retraction announcement was present on the journal homepage and/or in the PDF.

Results A total of 114 retracted articles in Korean medical journals were found using the KoreaMed database. On the journal homepage, retraction announcements were present for 47 of the 114 retracted articles (41.2%). Six articles (5.3%) contained a retraction announcement in the PDF. Of these 6 articles, 5 also had indications in the HTML file. There was no indication of retraction in 66 articles. Among the 114 retracted articles, 39 were cited in the Web of Science Core Collection (mean [SD] number of citations, 4.2 [4.4]; median number of citations, 2 [range, 1–19]), including a mean (SD) number of post-retraction citations of 2.4 (3.5) (median number of citations, 1 [range, 0–18]), and 41 were cited in Scopus (mean [SD] number of citations, 4.3 [3.9]; median number of citations, 3 [range, 1–18]), including a mean (SD) number of post-retraction citations of 2.6 (2.8) (median number of citations, 1 [range, 0–11]). In Web of Science, of the 134 total citations, 93 (69.4%) occurred 1 year after the retraction. In Scopus, of the 169 citations, 107 (63.3%) were post-retraction citations (Table 47). Whether the retraction announcement was made in the PDF or on the homepage did not influence post-retraction citations.

Conclusions Post-retraction citations are very common in Korean medical journals indexed in KoreaMed. The exact

Table 47. Frequency of Citations of Retracted Articles in Korean Medical Journals Obtained From the Web of Science Core Collection and Scopus Databases

Retracted Articles	Web of Science Core Collection		Scopus	
	No. of Articles	No. of Citations	No. of Articles	No. of Citations
Cited articles	39	134	41	169
Postretraction citation ^a	28	93	31	107
Non-postretraction citation	11	41	10	62
Noncited articles	75	0	73	0
Total	114	134	114	169

^aOne year after the retraction, excluding retraction-related citations.

reason is difficult to understand, and efforts should be made to identify the cause and to correct it.

¹Department of Parasitology and Institute of Medical Education, College of Medicine, Hallym University, Chuncheon, Korea; ²Medical Library, Hanyang University Guri Hospital, Guri, Korea; ³InfoLumi Co, Seongnam, Korea; ⁴Department of Family Medicine, Gangdong Sacred Heart Hospital, College of Medicine, Hallym University, Seoul, Korea, hallymfmm@gmail.com

Conflict of Interest Disclosures: H.-M. Cho is employed by InfoLumi Co. No other conflicts were reported.

Funding/Support: InfoLumi Co provided support in the form of a salary for H.-M. Cho but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

TUESDAY

Assessment of a Standardized Tool to Identify Deceptive Journals

Kathleen Berryman,¹ Sheree Crosby,¹ Lacey Earle,¹ Lucas Toutloff²

Objective Cabell’s provides academic journal analytics to the scientific research community. Our objective is to develop an unbiased, transparent, and effective tool for identifying deceptive academic journals by analyzing specific behavioral indicators.

Design For this study, we defined “deceptive” as intentional misrepresentation of facts or failure to provide implied services. Our first step was to identify behaviors common to journals recognized as deceptive. We examined 261 randomly selected journals from publishers that Jeffrey Beall identified as predatory in his widely accepted list of predatory publishers. We also referenced journals removed from the Directory of Open Access Journals in 2013. In addition, we systematically identified specific behaviors that contradicted industry standards and best practices. This process led us to 65 behavioral indicators of deception that were frequent or common in journals identified as deceptive. We then weighted each indicator and put it into 1 of 2 categories: indicators directly reflecting deception were weighted heavily and behaviors tending to coincide with deception were weighted lightly. The weights were carefully structured to

prevent bias against new or inexperienced journals. We then created a rubric and applied it to data collected about each journal. This produced a weighted score whose magnitude increased with the probability that a journal was engaging in deceptive behavior.

Results We applied the rubric to 2 different sets of journals: 1192 randomly selected journals from 57 publishers on Jeffrey Beall’s list and 100 journals from Cabell’s database. Of the 100 journals from Cabell’s, we randomly selected 50 journals from the top tier and bottom tier. Of the 1192 journals from Jeffrey Beall’s list, 1114 (93.5%) were flagged as deceptive by our methodology and 78 (6.5%) journals were not flagged as deceptive. Only 6 (10.5%) of the 57 total publishers examined had some journals flagged as deceptive and some not flagged. Our methodology did not identify any of the journals in the Cabell’s set as deceptive.

Conclusion The results led us to conclude that our methodology was effective in objectively identifying deceptive journals. The 78 journals on Jeffrey Beall’s list not identified by our methodology indicated the potential need for additional factors or tighter tolerances. However, any tightening of tolerances must not identify legitimate journals as deceptive. Further study into the methodology could be conducted to see if it identifies false positives from a set of journals that Jeffrey Beall evaluated but did not indicate as predatory.

¹Cabell’s International, Beaumont, TX, USA, kathleen.berryman@cabells.com

Conflict of Interest Disclosures: Cabell’s International intends to patent the methodology assessed in this paper and use it to create a list of deceptive journals that will be available by subscription.

Funding/Support: There was no external funding for this study. Contributions of staff time and resources came from Cabell’s International.

MONDAY

Assessment of the Prevalence of Integrity Issues in Submitted Manuscripts

Damian Pattinson,¹ Chrissy Prater¹

Objective There has been much discussion about the increase in articles with serious integrity breaches that are submitted to science journals. Anecdotaly journals are handling a growing number of these breaches, but their prevalence and distribution is unknown in part because most articles are rejected before publication.

Design Research Square is a company that works with publishers to perform detailed editorial checks to catch basic integrity issues before peer review. Articles that pass the checks that are performed by Research Square staff members are awarded a “badge” to show that they meet high standards of integrity. Each failed check is recorded using a standardized checklist, allowing analysis of the prevalence of various integrity issues, including plagiarism, figure manipulation, undisclosed competing interests, problems

Table 48. Categories of Integrity Concerns

Issue	Articles in Which a Check is Applicable, No. (%)	Articles Failed, No. (%) ^a
Inability to verify author identities	2892 (100)	1001 (35)
Plagiarism	2892 (100)	648 (22)
Missing funding statement	2892 (100)	434 (15)
Undisclosed conflicts of interest	2892 (100)	253 (9)
Missing human study participant approval	2060 (71)	880 (43)
Image manipulation	459 (16)	100 (22)
Missing animal ethics approval	153 (5)	25 (16)

^aThis percentage is calculated as the number of failed articles of the total number for which that check was applicable.

surrounding permissions, and lack of ethical approval or funding statements.

Results Of 2892 checklists from manuscripts submitted to 2 mid-sized open access life science journals (Impact Factors of 3 and 2, respectively), 628 (22%) passed all checks. Of the article with applicable checklists, 22% of submissions were returned for rewriting because of plagiarism and 22% of applicable manuscripts failed figure checks, with the top 3 reasons being improper manipulation, duplication, and poor quality. Other common issues included an inability to verify author identities (35%) and missing statements of approval for human participation in research (71%) (**Table 48**). While many issues could be solved with author queries, 161 submissions (5%) contained serious flaws that placed the veracity of the article in doubt, including defamatory content, extensive plagiarism from published works, and suspicion of fabrication. Results varied by journal.

Conclusions Duplicate publication, figure concerns, inability to verify authors, and missing human participant approval were the most common of integrity concerns at 2 life science journals. Additional analyses of attributes of articles that failed integrity checks and their authors might yield insights into how to improve the detection of such issues in the future, especially when applying the checklist to journals in a broader range of subject areas.

¹Research Square, Durham, NC, USA, damian.pattinson@researchsquare.com

Conflict of Interest Disclosures: All authors are employed by Research Square, a private company that provides language and editorial services for publishers and authors.

Funding/Support: None reported.

Peer Review

TUESDAY

A Comparison of Reviewer Contribution Distributions in Publons

Andrew Preston,¹ Tom Culley¹

Objective A 2015 study by Kovanis et al used an early Publons data set to develop a mathematical model that showed an imbalance in the distribution of peer review effort across the biomedical research community. They concluded that 20% of researchers performed between 69% and 94% of reviews. Their study focused on biomedical literature, but the Publons data set was not restricted to that specific research field. Publons is a platform that allows peer reviewers to track, verify, and claim credit for their peer review contributions across the world's journals. Since 2015, the Publons platform has grown significantly. More than 1350 journals from many disciplines now integrate directly with the system, giving a broader sampling of peer reviewer contributions. In this study we present an updated data set to report the distribution of reviewer workload in the biomedical sciences and other research fields, seeking to confirm whether the heavy reviewing workload is borne by a small proportion of reviewers across all research fields.

Design We used empirical distribution of prepublication review contributions recorded for the 2015 calendar year by researchers on Publons and tagged by discipline (as defined by the All Science Journal Classification provided by SCOPUS and country. Researchers more likely to have incomplete review records for any given year were filtered out (by only including researchers who signed up ≥ 3 months after the completion of that year but had added their reviews from the previous year) as a means of controlling for incomplete Publons records and bias.

Results Overall, there were 51,482 reviews completed by 35,248 reviewers in Publons in 2015 (**Table 49**). Of these, 23,600 reviewers (67%) were excluded because they were likely to have incomplete records. This left 11,648 reviewers from 26 research fields and 126 countries. There was a more even distribution of reviews performed by reviewers than that previously reported by Kovanis et al (20% of reviewers performed 59% of the reviews compared with the 69% to 94% reported by Kovanis et al). Reviews in biomedical fields ($n = 4583$) were more evenly distributed, with 20% of reviewers performing 51% of the reviews compared with the 41,719 reviews in the nonbiomedical fields (20% of reviewers performed 58% of reviews). Analysis by country is ongoing and will be presented.

Conclusions If the distribution of peer review records added by researchers active as reviewers on Publons is representative of the distribution of global prepublication peer review effort, then the peer review burden appears more evenly spread than previously suggested by Kovanis et al. The distribution is still skewed, with a small proportion of researchers shouldering much of the reviewing load, but researchers in the biomedical field seem to shoulder a more equitable load than those in other fields. Nevertheless, funders, research institutions, and publishers may want to recognize overburdened reviewers or encourage other researchers to contribute more to ensure that the quality and supply of peer review keeps up with the increasing number of article submissions.

Table 49. Proportion of Reviews Performed by 20% and 50% of Reviewers in 2015 Across Legacy and Updated Publons Databases

Reviewers, %	Reviews Performed, No. (%)			
	Kovanis Report, % ^a	All Publons Fields (51,482 Reviews)	Publons Biomedical Fields ^b (4583 Reviews)	Publons Nonbiomedical Fields ^c (41,719 Reviews)
20	69-94	30,374 (59)	2337 (51)	24,197 (58)
50	88-100	43,759 (85)	3437 (75)	35,043 (84)

^aThe report by Kovanis et al (<https://doi.org/10.1371/journal.pone.0166387>) used a combination of Publons and MEDLINE data to estimate the global burden of journal review in the biomedical literature.

^bBiomedical fields include: biochemistry, genetics, and molecular biology; immunology and microbiology; medicine; neuroscience; and pharmacology, toxicology, and pharmaceuticals.

^cReviews from 22 other nonbiomedical fields (according to the All Science Journal Classification provided by SCOPUS [http://ebbrp.elsevier.com/pdf/Scopus_Custom_Data_Documentation_v4.pdf]) were not included.

^dPublons, London, UK, andrew@publons.com

Conflict of Interest Disclosures: Andrew Preston is the managing director and a cofounder of Publons; Tom Culley is the marketing director. Publons is a limited liability company with offices in London, UK, and Wellington, New Zealand, and was acquired by Clarivate Analytics in 2017.

MONDAY

A Pilot Study of Online Training of Patient Stakeholders Aimed at Improving Knowledge and Skills to Complete Peer Review of PCORI Draft Final Research Reports

Karen B. Eden,¹ Ilya Ivlev,¹ Amy Forester,² Camber Hansen-Karr,¹ Ed Reid,¹ Lauren Saxton,¹ Kelly Vander Ley,¹ Mark Helfand¹

Objective Many patient, caregiver, and patient advocate reviewers for the Patient-Centered Outcomes Research Institute (PCORI) have not been formally trained in peer review. This pilot study was designed to develop new patient-oriented web-based PCORI peer review training and evaluate the effect of the new training on participants' knowledge and self-efficacy to serve as a peer reviewer.

Design In September 2016, we conducted a qualitative study to assess complexity of the task of reviewing a PCORI Draft Final Research Report (DFRR) and to identify training needs. We then conducted a pilot before-after study of early users of a PCORI patient peer review web-based training between February 2017 and May 2017. The training site, hosted by Sakai, Oregon Health & Science University's learning management system, included videos on PCORI and the peer review process, a learning activity based on a DFRR annotated by editors, a writing activity, sample reviewer comments that were more helpful and less helpful, and online resources. Using site analytics within Sakai, access to the 6 training components was tracked. In this pilot study, we analyzed pretraining and posttraining knowledge (15 questions) and skills (6 questions) using McNemar tests; self-efficacy and attitude toward peer review were analyzed using Wilcoxon tests.

Results Fourteen participants were interviewed about their experience reviewing a DFRR and 33 reviewers (14 patients, 18 patient advocates, and 1 caregiver) completed the pilot cross-sectional, before-after phase (33 of 52 [63%] completed pretraining and posttraining surveys and reported completing the training). Twelve of 14 interviewed participants reported having difficulties understanding from what point of view they needed to assess a technically written DFRR and reported a need for training in peer review. After the training was developed, 28 of 33 participants in the before-after study accessed 4 or more of the 6 training components. The percentage of correct answers, across participants, increased significantly after training for the knowledge questions (343 [71.5%] vs 441 [91.9%] of 480 answers; $P < .001$) (Table 50) and skills questions (143 [74.5%] vs 170 [88.5%] of 192 answers; $P < .001$). The number of questions answered as unsure significantly decreased for both knowledge questions (62 [12.9%] vs 6 [1.3%] of 480 questions; $P < .001$) and skills questions (19 [9.9%] vs 3 [1.6%] of 192 questions; $P < .001$). Importantly, the training increased confidence in patient stakeholders that they can complete a high-quality peer review ($Z = -2.69$; $P = .007$). All 33 participants reported that the training enhanced their knowledge and that they would recommend training to other patient or caregiver reviewers.

Table 50. Pretraining and Posttraining Data

Outcome	Pretraining, No. (%)	Posttraining, No. (%)	P Value
Knowledge questions among respondents who answered before and after training (n = 32)			
No. of correct answers/total No. of answers (%)	343/480 (71.5)	441/480 (91.9)	<.001 ^a
No. of unsure answers/total No. of answers (%)	62/480 (12.9)	6/480 (1.3)	<.001 ^a
Skills questions among respondents who answered before and after training			
No. of correct answers/total No. of answers (%)	143/192 (74.5)	170/192 (88.5)	<.001 ^a
No. of unsure answers/total No. of answers (%)	19/192 (9.9)	3/192 (1.6)	<.001 ^a
Response to statement "I am confident I can complete a high-quality peer review of DFRR for PCORI," to assess self-efficacy (n = 33)			
Strongly agree	13 (39.4)	22 (66.6)	.007 ^b
Agree	13 (39.4)	9 (27.3)	
Undecided	6 (18.2)	2 (6.1)	
Disagree	0	0	
Strongly disagree	1 (3.0)	0	
Response to statement "I am excited to provide a peer review of a DFRR for PCORI," to assess attitude (n = 33)			
Strongly agree	21 (63.7)	24 (72.7)	.09 ^b
Agree	10 (30.3)	8 (24.2)	
Undecided	1 (3.0)	1 (3.0)	
Disagree	1 (3.0)	0	
Strongly disagree	0	0	

Abbreviations: DFRR, Draft Final Research Report; Patient-Centered Outcomes Research Institute.

^aMcNemar test.

^bWilcoxon test.

Conclusions Completion of an online peer review training program shows promise in increasing knowledge and skills of reviewers. In the future, we plan to evaluate the dose of online training with the quality of actual peer reviews by patients and caregivers.

¹Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, OR, USA, edenk@ohsu.edu; ²Teaching and Learning Center, Oregon Health & Science University, Portland, OR, USA

Conflict of Interest Disclosures: None reported.

Funding/Support: This study was conducted under a contract to provide peer review services for the PCORI DFRRs. Karen B. Eden, Amy Forester, Camber Hansen-Karr, Ed Reid, Lauren Saxton, Kelly Vander Ley, and Mark Helfand receive support from PCORI for developing the patient peer review training program and reviewing reports. Ilya Ivlev is a postdoctoral fellow supported by grant T15LM00708822 from the National Library of Medicine.

Role of the Funder/Sponsor: PCORI was not involved in conducting this research study, but PCORI members were part of the team that prepared materials (videos, review form) for the patient and caregiver training. PCORI reviewed and approved the abstract for submission.

Acknowledgments: We thank Kira Lesley for her timely help in recruiting patients, patient advocates, and caregivers.

MONDAY

Identification of Motivations for Peer Reviewers to Perform Prepublication Review of Manuscripts: A Systematic Review

Mersiha Mahmić-Kaknjo,¹ Mario Malički,² Ana Utrobičić,³ Dario Sambunjak,⁴ Ana Marušić²

Objective To identify and synthesize studies regarding motivation for prepublication peer review of manuscripts.

Design A Systematic review of studies indexed in MEDLINE, Web of Science (WoS) and Scopus was carried out. Literature search was performed in February, 2016 with no language or time limitations. A total of 3585 records remained after deduplication. Initial screening of titles and abstracts was conducted by 2 independent reviewers. For all records without indexed abstracts, full text was obtained. Bibliographies of selected studies are still to be examined to identify additional relevant studies. Qualitative studies were assessed using the Consolidated Criteria for Reporting Qualitative Research (COREQ) and surveys by using the Good Practice in the Conduct and Reporting of Survey Research Checklist.

Results Of 3585 records, 315 were related to peer review, but only 14 explored motivations for prepublication peer review of manuscripts, of which 4 were agent-based models (simulations) dealing with peer review incentives, 4 were qualitative studies of reviewers or editors (total 94 participants), 3 were surveys (total 2308 respondents, participation rates 62%, 63%, and not listed), and 3 were theoretical papers on new indices or incentives that would improve the motivation of reviewers and quality of their

reviews. Both surveys and qualitative studies reported the following most common incentives to peer review: contribute to the community/scientific field, reciprocity, keep up to date on current research, improve manuscript quality, acquire new skills and experience, and career advancement. The most common disincentives were: lack of time, poor quality of manuscript or journals, and lack of formal recognition of performed work.

Conclusions Studies on motivation for performing peer review are rare. Most reported incentives for conducting reviews were contributing to the community and keeping up to date with new studies, with lack of time being the most common reason for refusing to review. After checking bibliographies of selected studies for additional studies we will attempt a synthesis of the results.

¹Department of Clinical Pharmacology, Cantonal Hospital, Zenica, Bosnia and Herzegovina; mmahmickaknjo@gmail.com;

²Cochrane Croatia and Department of Research in Biomedicine and Health, University of Split, School of Medicine, Split, Croatia;

³Cochrane Croatia and Central Medical Library, University of Split School of Medicine, Split, Croatia; ⁴Catholic University of Croatia, Department of Nursing, Center for Evidence-Based Medicine and Health Care, Zagreb, Croatia

Conflict of Interest Disclosures: Dr Marušić is a Peer Review Congress Advisory Board Member but was not involved in the review or decision for this abstract. No other conflicts were reported.

Funding/Support: This research was funded by COST Action TD1306 New frontiers of peer review (PEERE).

Role of the Funder/Sponsor: The funder had no role in the design, execution, interpretation, or writing up of the study.

TUESDAY

Publishing Peer Review Reports Alongside Articles With Separate DOIs: A Pilot Study of 5 Journals in Different Scientific Disciplines

Bahar Mehmani¹

Objective This study assessed reviewers', editors', and authors' views about the publication of peer review reports (signed or anonymous depending on the reviewer's consent) alongside articles with separate DOIs to give more credit to reviewers and add more transparency to the process.

Design Since February 2014, Elsevier has been publishing review reports of accepted manuscripts with separate DOIs on ScienceDirect from 5 journals in different scientific disciplines (*Annals of Medicine and Surgery*, *Agricultural and Forest Meteorology*, *Engineering Fracture Mechanics*, *International Journal of Surgery*, and *Journal of Hydrology*). Reviewers can sign their report or remain anonymous. Survey questionnaires were sent (by email without reminder) to 644 reviewers (204 who had agreed to review and 440 who had declined to review by August 2015), 40 editors, and 3774 authors to assess the usefulness of this exercise. Emails were sent with a link to questionnaires for authors, reviewers, and editors, with a mix of response

options (yes or no, multiple choice, Likert scale, and open answer).

Results Of the 204 reviewers who agreed to review, 40 (19.6%) responded to the survey. Of these 40 respondents, 38 (95.0%) indicated that publishing peer review reports did not influence their recommendation, and 39 (97.5%) indicated they would accept further review invitations from the journal. Of the 440 reviewers who declined to review, 100 (22.7%) responded to the survey. Of these 100 respondents, 91 (91.0%) indicated that publication of the review report was not a reason for declining, 68 (68.0%) indicated lack of time as the primary reason, 24 (24.0%) indicated personal reasons, and 23 (23.0%) indicated a mismatch with areas of expertise. Sixteen of 40 editors responded to the survey (40.0%). Five editors indicated that the pilot study made it more difficult for them to find a reviewer, 2 indicated that the pilot study did not have any influence, and 9 were undecided. Ten journal editors said that since the pilot study, they noticed that the reports were more in depth and constructive for authors, 5 did not notice any difference, and 1 did not reply to this question. Six editors mentioned that they used the published review reports of their journal as examples of the reviewing process for younger reviewers. Of the 3774 authors whose manuscripts were accepted by the 5 journals and who were sent the survey, 501 (13.3%) responded (**Table 51**). The percentage of authors who indicated that they like the publication of the peer review reports ranged from 48.7% (*Engineering Fracture Mechanics*) to 60.5% (*Annals of Medicine and Surgery*), and 51.6% to 63.9% of authors indicated that this policy would not influence their decision where to publish.

Conclusions Although the findings are limited by the low response rates, the communities served by the journals in this study seem to be open to the practice of publishing peer review reports. Based on the result of this pilot study, Elsevier will make it possible for reviewers of its other journals from similar research areas to publish their review reports.

¹Elsevier, Amsterdam, the Netherlands; b.mehmani@elsevier.com

Conflict of Interest Disclosures: None reported.

Funding/Support: This study was funded by Elsevier.

Additional Contributions: Dr Mehmani acknowledges 5 pilot journal editors for supporting the pilot.

TUESDAY

Peer Review in Computer Science Conferences Published by Springer

Mario Malički,¹ Martin Mihajlov,² Aliaksandr Birukou,³ Volha Bryl⁴

Objective To describe the types of peer review, number of reviewers, use of external reviewers, acceptance rates, and submission systems associated with computer science conference proceedings published by Springer.

Table 51. Survey Responses of Authors of Accepted Articles Regarding a Change in Policy to Publish Peer Review Reports Alongside Articles

Question and Response	Respondents, No. (%)				
	Agricultural and Forest Meteorology	Journal of Hydrology	Annals of Medicine and Surgery	International Journal of Surgery	Engineering Fracture Mechanics
Does this change in policy influence your decision to publish in the journal?					
Respondents, No.	172	62	40	105	122
This will make it more likely for me to publish in the journal.	47 (27.3)	24 (38.7)	14 (35.0)	39 (37.1)	37 (30.3)
This will make it less likely for me to publish in the journal.	15 (8.7)	6 (9.7)	5 (12.5)	10 (9.5%)	11 (9.0)
It will have no influence on my decision where to publish.	110 (63.9)	32 (51.6)	21 (52.5)	56 (53.3)	74 (60.7)
How did you like publication of peer review reports next to your article?					
Respondents, No.	156	51	38	100	117
I liked it very much	21 (13.5)	7 (13.7)	8 (21.1)	14 (14.0)	18 (15.4)
I liked it	67 (42.9)	22 (43.1)	15 (39.5)	43 (43.0)	39 (33.3)
I don't care	53 (34.0)	13 (25.5)	10 (26.3)	36 (36.0)	42 (35.9)
I disliked it	11 (7.1)	7 (13.7)	3 (7.9)	4 (4.0)	9 (7.7)
I didn't like it at all	4 (2.6)	2 (3.9)	2 (5.3)	3 (3.0)	9 (7.7)

Design We used the Springer online delivery platform to identify computer science conference proceedings published by Springer between 1973 and 2017. Proceeding prefaces list published articles and detail the process used for their selection; we batch-downloaded all prefaces using JDownloader software, converted the prefaces to text using UNIPdf or Adobe Acrobat Pro, and extracted peer review information using a combination of regular expression matching with Perl and manual data curation.

Results Of 7710 processed conference prefaces, we identified 1657 unique conferences that had been held a median of 3 (range, 1–39) times. Prefaces contained information on number of submissions sent for review (n=5021 [65%]), accepted full articles (n=4890 [63%]), short articles (n=940 [12%]), posters (n=711 [9%]), peer review type (n=561 [7%]), number of reviewers (n=2962 [38%]), use of external reviewers (n=2716 [35%]), and submission systems (n=1392 [18%]). Acceptance rates for full articles ranged from 3% to 93%, with a time-weighted median of 37% (95% CI, 36%–39%). Acceptance rates tended to decrease per number of times the conference was held (**Figure 10**). Conferences used a median 3 reviewers per article (range, 2–11) and most commonly used double-blind peer review (315/561 [56%]), although a small proportion (32/237 [14%]) changed their type of peer review over time (eg, from single to double blind). The most common submission systems used were EasyChair, CyberChair, and iChair.

Conclusions Computer science conferences decrease their acceptance rates over time. The reasons were not explored in this study but may be due to an increase in reputation of the conferences and in number of submissions over time.

¹Cochrane Croatia, Department of Research in Biomedicine and Health, Department of Medical Humanities, University of Split School of Medicine, Split, Croatia; ²Faculty of Economics, Ss Cyril and Methodius University, Skopje, Macedonia; ³Springer Verlag

GmbH, Heidelberg, Germany, aliaksandr.birukou@springer.com; ⁴Springer Nature, Heidelberg, Germany

Conflict of Interest Disclosures: None reported.

Funding/Support: This research was funded by COST Action TD1306, New Frontiers of Peer Review (PEERE). PEERE awarded Malički Mario and Martin Mihajlov the Short Term Scientific Mission to extract the peer review data and citation data and will cover the publication costs of the manuscript. The initial prototype of the Springer Linked Open Data pilot for computer science conferences was developed by Net Wise in cooperation with University of Mannheim and was partially supported by the LOD 2 EU FP7 project.

MONDAY

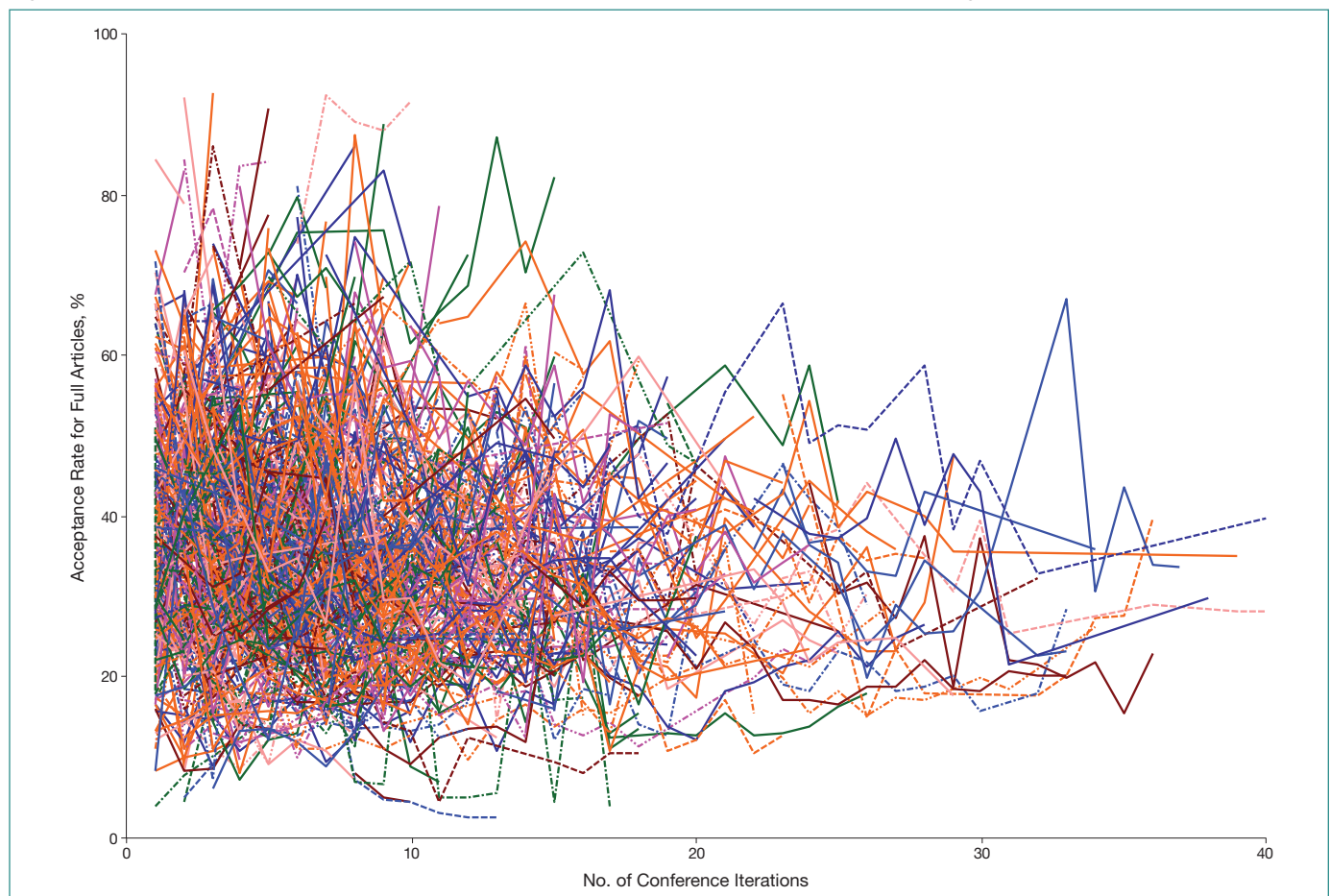
An Analysis of Peer Review Cases Brought to the Committee on Publication Ethics (COPE) From 1997 to 2016

Elizabeth C. Moylan,¹ Virginia Barbour,² Linda Gough,³ Charon A. Pierson,⁴ Deborah Poff,⁴ Natalie Ridgeway,³ Michael Wise,⁶ Adrian Ziderman⁷

Objective The Committee on Publication Ethics (COPE) holds a quarterly forum where editors from its membership can raise cases in publication ethics for discussion and advice. All of the forum cases from 1997 onwards have been entered into a searchable database. This database contains over 500 cases together with the advice given by COPE. More recent cases also include follow-up information and outcomes, providing a valuable resource for users. Our aim was to analyse the cases that are related to peer review in the context of the other categories of cases brought to COPE. This work builds on the COPE Case Taxonomy, specifically reporting on trends within peer review to inform future guidance.

Design Overall, 43 cases from the “peer review” classification category from 1997 to 2016 were suitable for further analysis. The specific issue within peer review was noted, together with

Figure 10. Time Series of Acceptance Rates of Computer Science Conferences Published by Springer (n=985)



the date the specific issue arose and who was responsible at that point (author, peer reviewer, editor). The broad subject area of the journal, as well as the recommendation for the journal, was also recorded.

Results The number of cases brought to COPE that involve peer review have declined in recent years, as have all other previously high-frequency categories apart from cases relating to authorship, which remain consistently high. Within the peer review cases, 29 occurred in science disciplines and 3 occurred in the arts and social sciences. In 11 cases the subject area was not known. Most issues arose during the peer review process (32 cases) with breaches in confidentiality and concerns about the editorial process being the most common reasons why cases are brought to COPE. However, in recent years more complex issues relating to bias in peer review and compromised peer review (due to fraud) have occurred. Although each case has unique outcomes and specific advice, a common finding was the recommendation to revise journal guidelines and policy (**Table 52**).

Conclusions The decline in incidence of cases brought to COPE in previously high-frequency categories (except for authorship) may reflect the use of the COPE cases database which facilitates users learning from related cases. It may also reflect a growing awareness of relevant COPE guidance in this area (eg, ethical guidelines for peer reviewers). However, there is potential for further guidance in peer review,

particularly with respect to maintaining confidentiality and best practice in editorial processes, as well as for handling cases involving more than one journal.

¹BioMed Central, London, UK, elizabeth.moylan@biomedcentral.com; ²Queensland University of Technology, Queensland, Australia; ³COPE, UK; ⁴American Association of Nurse Practitioners; ⁵Pacific Coast University for Workplace Health Sciences, British Columbia, Canada; ⁶University of Western Australia, Perth, Australia; ⁷Bar-Ilan University, Ramat Gan, Israel

Conflict of Interest Disclosures: Dr Moylan is a COPE council member (unpaid) and senior editor (research integrity) at BioMed Central. She is an editorial board member for *Research Integrity and Peer Review* and a member of the advisory board for EnTIRE (an EU funded proposal for mapping the research ethics and research integrity framework). Dr Barbour is the immediate past chair of COPE (unpaid). She is also a member of the Peer Review Congress Advisory Board but was not involved in the review or decision of this abstract. Ms Gough is COPE's administrator and is a paid employee of COPE. Dr Pierson is secretary and interim treasurer of COPE (unpaid), as well as the editor in chief for the *Journal of the American Association of Nurse Practitioners* (paid position by the American Association of Nurse Practitioners). Dr Poff is vice-chair and chair elect of COPE (unpaid). She is the cofounder and editor of the *Journal of Business Ethics* and editor in chief of the *Journal of Academic Ethics*, as well as the editor of *Business Ethics in Canada* and the section editor on Business and Economic Ethics of *Encyclopedia of Applied Ethics*. Ms Ridgeway is COPE's executive officer and is a paid employee of COPE. Dr Wise is a COPE council member (unpaid). He receives research support from Australian governmental bodies and

Table 52. Frequency of Various Issues Within Peer Review Cases Brought to Committee on Publication Ethics, 1997 to 2016

Specific Issue Within Peer Review	Frequency	Consistent or Recent Issue	Years Issue Occurred
Bias in peer review	2	Recent	2015, 2016
Breach in confidentiality	13	Consistent	1997, 2005, 2007, 2008, 2010, 2012, 2013, 2014, 2016
Conflicts of interest	6	Consistent	1999, 2005, 2009, 2010, 2012
Compromised peer review	3	Recent	2011, 2012
Consent for publication	1	Recent	2012
Editorial decision making	5	Consistent	2000, 2003, 2004, 2011
Editorial policy	1	Recent	2010
Editorial process	9	Consistent	2002, 2007, 2008, 2009, 2011, 2012, 2013
Legal concerns	1	Rare	2003
Member conduct	1	Rare	1997
Models of peer review	1	Recent	2016

charitable trusts. The Immunisation Alliance of Western Australia (of which he is a member) has received donations and/or funding from both individuals, the Government of Western Australia and from companies including GlaxoSmithKline, CSL, and Sanofi Pasteur. Dr Ziderman is a trustee of COPE (unpaid). He is editor in chief of *International Journal of Manpower*, and president of the International Network of Business Management Journals. No other conflicts are reported.

MONDAY

Modeling the Effects of Jointly Implemented Peer Review Systems on Scientific Publication

Michail Kovanis,^{1,2} Philippe Ravaud,^{1,5} Raphael Porcher¹⁻³

Objective To estimate the effects of joint implementation of different peer review (PR) systems on overall PR efficiency, reviewer effort, and scientific dissemination.

Design We used a previously developed model of the scientific publication system with data inputs from surveys of researchers (conducted by us and Elsevier) and from journals (Journal Citation Reports 2013). In the model, researchers were characterized by their resources and scientific level, papers by their scientific value, and journals by their reputation and acceptance/rejection thresholds. Papers could be reviewed and revised, then accepted or rejected, then resubmitted or abandoned if rejected. We considered the conventional PR system and 5 alternatives: (1) up to 1 round of reviews and revisions allowed (re-review opt-out), (2) rejected papers resubmitted with their past reviews to another journal (portable PR), (3) rejected papers resubmitted with their past reviews to journals of lower reputation within 1 publisher network (cascade PR), (4)

articles available online on submission and editors considered both invited reviews and comments from the community (crowdsourcing PR), and (5) a similar system without online comments (immediate publication). We modeled the effects of joint implementation of the 5 systems (**Figure 11**) on PR (measured as the separation between distributions of scientific value of published and unpublished papers [Hellinger distance] and as the mean relative increase in scientific value), on reviewer effort (measured as the total time spent in PR), and on scientific dissemination (measured as the median time from submission to final decision, and the average weekly release of scientific information [paper scientific value × journal reputation]) varied by journal adoption rates.

Results Compared with conventional PR, scenarios in which review-sharing systems (alternative 2 or 3) were dominant (≥50% adoption) had the greatest impact on scientific publication: the separation of the distributions of scientific value was unchanged, the mean changes in papers’ scientific value increased by 1.3% to 13.8% (range of means across scenarios), the total time devoted to PR decreased by 33.5% to 64.3%, the median time to final decision decreased by a relative 31.4% to 47.3% and the release of scientific information increased by 3.1% to 36.5%. The scenario with equal (16.7%) adoption of all alternative systems was almost as beneficial.

Conclusions In this simulation of joint implementation of different PR systems, review-sharing systems seemed most promising at increasing PR efficiency and decreasing PR effort, and may be further tested in real-world trials.

¹INSERM U1153, Paris, France, michail.kovanis@aphp.fr; ²Université Paris Descartes–Sorbonne Paris cité, Paris, France; ³Assistance Publique–Hôpitaux de Paris, Hôpital Hôtel–Dieu, Centre d’Epidémiologie Clinique, Paris, France; ⁴Cochrane France, Paris, France; ⁵Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA

Conflict of Interest Disclosures: None reported.

TUESDAY

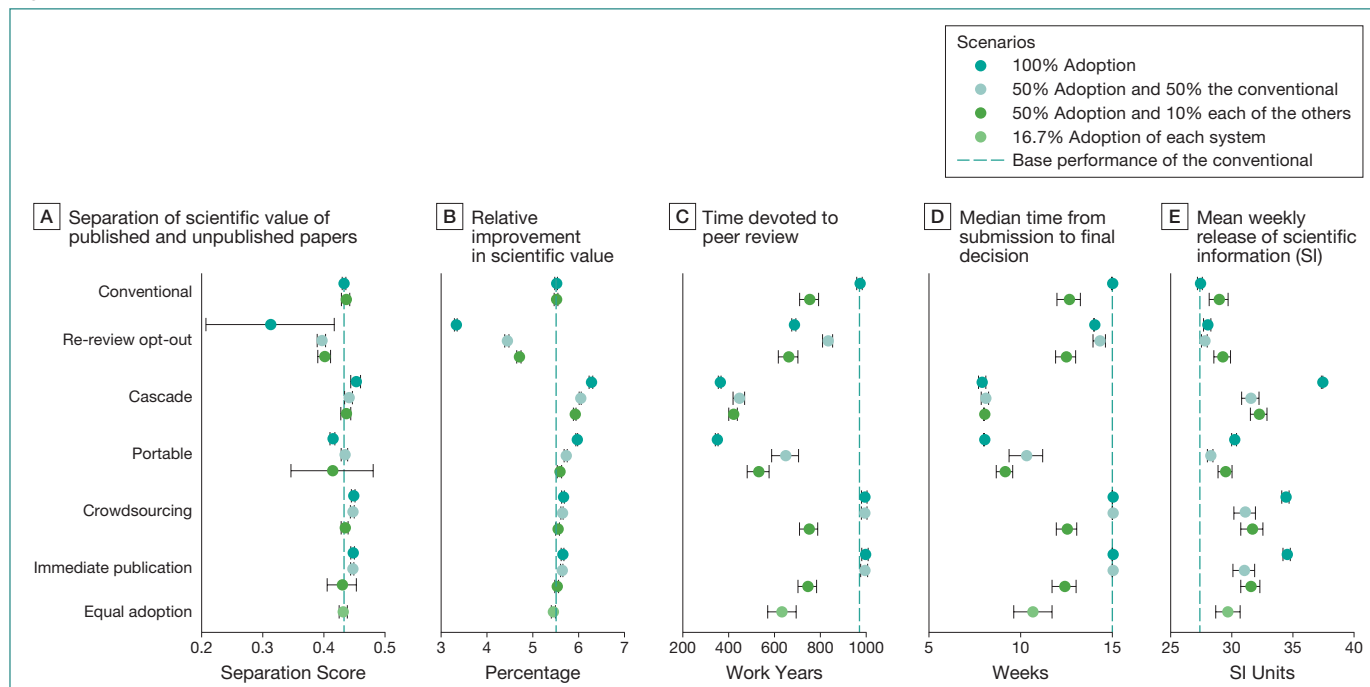
Linguistic Features in Peer Reviewer Reports: How Peer Reviewers Communicate Their Recommendations

Ketevan Glonti,^{1,2} Darko Hren,¹ Simon Carter,³ Sara Schroter⁴

Objective Many biomedical journals ask peer reviewers to provide a recommendation on the manuscript under review: whether to accept with no revision; minor revision; major revision, or to reject the manuscript. Some editors use these reviewer recommendations to help their editorial decisions. The content and linguistic features of peer reviewer reports may vary depending on the reviewer’s recommendation. Our aim was to identify trends in the style of language employed by peer reviewers when providing a recommendation.

Design This is a retrospective analysis of peer reviewer reports collected during a previously reported single-blind

Figure 11. Model Results for Peer Review Systems



randomized controlled trial carried out by *The BMJ*. That trial explored the effects of providing training to reviewers on the quality of peer reviewer reports of 3 manuscripts. For this study, we analyzed a random sample of 440 of 1372 reviewer reports in the control arm of the study. Of these, 330 were recommended for rejection and 110 were recommended for acceptance (8 were recommended for acceptance with no revision, 38 for acceptance with minor revision, 42 for acceptance with major revision, and 22 other recommendations. Peer reviewer reports were analyzed using automatic textual analysis software Linguistic Inquiry and Word Count (version 1.3.1, LIWC2015) to capture the rhetorical strategies (eg, analytic language) deployed in the text in 4 language variables: analytic, clout, authentic, tone. These LIWC scores for individual reviews were imported into an SPSS database (version 24, IBM) where descriptive statistics (means, standard deviations, and 95% CIs for means) were calculated and statistical tests performed using independent sample *t* tests.

Results Table 53 summarizes key differences in language strategies used in peer reviewer reports that recommend acceptance or rejection of the manuscript being reviewed. There was no difference in reviewer reports recommending rejection vs acceptance in their analytic nature. Reviewer reports recommending rejection were significantly more "authentic" (ie, had more frequent use of terms that indicate honesty and personal disclosure) and scored lower in terms of "tone" (ie emotional tone, where low scores indicate regret or hostility). Reviewer reports recommending rejection also scored lower in terms of "clout" (for which, high scores suggest that a review emerges from a perspective of high expertise and confidence on the part of the reviewer, and low scores reflect a more tentative or humble language style).

Table 53. Summary Language Variables Per Recommendation

Variable	Decision	Count	Mean	SD	95% CI for Mean
Analytic	not reject	110	85.2	11.5	83.0-87.3
	reject	330	86.2	9.0	85.2-87.2
Clout	not reject	110	41.6	10.0	39.7-43.5
	reject ^a	330	39.1	10.1	38.0-40.2
Authentic	not reject	110	17.7	12.4	15.4-20.0
	reject ^b	330	22.4	12.9	21.1-23.8
Tone	not reject	110	61.1	22.4	56.9-65.3
	reject ^b	330	45.5	19.6	43.4-47.6

^aStatistically significant difference from "not reject" (independent sample *t* test, *P* = .03)

^bStatistically significant difference from "not reject" (independent sample *t* test, *P* ≤ .001)

Conclusions Many journals in the biomedical field request that peer reviewers make a recommendation on the fate of a submitted manuscript. Greater awareness and understanding of the type of feedback and linguistic features employed by peer reviewers to communicate their recommendation may enable editors to develop strategies for making better informed editorial decisions.

¹Department of Psychology, School of Humanities and Social Sciences, University of Split, Split, Croatia, kglonti@unist.hr,² Université Paris Descartes, Sorbonne Paris Cité, Paris, France,³ Department of Sociology, The Open University, Milton Keynes, UK,⁴ BMJ, London, UK

Conflict of Interest Disclosures: Ketevan Glonti and Darko Hren were supported by a grant by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 676207. Sara Schroter is an employee of *The BMJ* and regularly undertakes research into the publishing process. Simon Carter reported no conflicts of interest.

Assessment of Regional Diversity of Reviewers in Journals Published in Medicine and Agricultural and Biological Sciences

Thomas Gaston,¹ Pippa Smart²

Objective This research investigated regional diversity of reviewers to assess whether there were variations that could be attributed to the journal’s (1) discipline, (2) size, (3) rank, (4) editor in chief (EIC) location, and (5) author location. The hypotheses were that the EIC and the author locations will affect reviewer selection and that lower-ranking journals will look more widely for reviewers.

Design The research used 2016 ScholarOne data for all Wiley-owned journals in medicine (n = 112) and the agricultural and biological sciences (n = 37). The EIC, reviewer, and author locations were determined by current institution, not country of origin. The journals were classified into large (>200 articles published each year), medium (100-200), and small (<100) publications; 132 journals had Impact Factors.

Results In all, 148 EICs, 110,053 reviewers, and 55,732 manuscripts were included in our analysis. The EICs were based in Asia (9 EICs), Europe (70), North America (65), Oceania (4). A correlation was found between EIC and reviewer locations. For each EIC region, the percentage of reviewers from the same region was higher than the overall mean value (Asia, +22%; Europe, +12%; North America, +13%; and Oceania, +4%) (**Table 54**). In addition, a preference was noted for reviewers from the same country as the EIC, although this rarely exceeded the preference for USA reviewers. For example, UK EICs invited 22% of their reviewers from individuals based in the UK but 27% from the United States; German EICs, 13% from Germany vs 26% from the United States. No similar correlations were found between the region of the EIC and whether invited reviewers were likely to accept the invitation. In all, 25.7% of authors came from Asia, but only 9.1% of reviewers. The global ratio of authors to reviewers was similar for both disciplines and all journal sizes. There was some evidence of reviewer bias: agreeing to review (54% vs 49%) and recommending acceptance (51% vs 46%) for authors in the same region. No significant correlations were noted between location of reviewer and other factors. Limitations of the study include the low number of Asian and high number of UK and USA

editors, which may have led to skewed data. In addition, Wiley is a USA-based publisher with a largely English-language output.

Conclusions The imbalance of author and reviewer location supports previous research. The journal EIC and author location are more influential on reviewer location than the journal profile, suggesting use of the EIC’s networks supported with a “local” review. Reviewers may be more positive if the author is from their region. There is little difference between the disciplines, although agricultural and biological science journals have a more global range of authors and reviewers.

¹Wiley, Oxford, UK, tgaston@wiley.com; ²PSP Consulting, Appleton, Oxfordshire, UK

Conflict of Interest Disclosures: Thomas Gaston is employed by Wiley.

Acknowledgments: We thank Sandy Garel and Katy Underhill (Oxford Brookes) for their assistance with data gathering and Kornelia Junge (Wiley) for her assistance with the data analysis.

Quality of Reporting

Characterizing Major Issues in ClinicalTrials.gov Results Submissions

Heather D. Dobbins,¹ Cassiah Cox,¹ Tony Tse,¹ Rebecca J. Williams,¹ Deborah A. Zarin¹

Objective The ClinicalTrials.gov results database provides public access to aggregate results data for nearly 27,000 clinical studies (as of June 2017). Results submissions must include all prespecified primary and secondary outcomes, all serious adverse events, and may not include subjective narratives. Results submissions are validated using 3 components: (1) structured data elements, (2) automated system-based checks, and (3) manual review by ClinicalTrials.gov staff. Since 2009, ClinicalTrials.gov has refined manual review criteria covering major issues that must be corrected or addressed as well as advisory issues; the goal is to avoid apparent errors, deficiencies, or inconsistencies and to ensure complete, sensible entries that can be understood by readers of the medical literature. We set out to characterize the type and frequency of major issues identified by manual review in results submissions.

Table 54. Region of Invited Reviewer by Region of Editor in Chief

Editor in Chief Region	No. (%) of Reviewers by Region ^a							
	Africa	Asia	Central/South America	Europe	Middle East	North America	Oceania	Total
Asia	44 (0.75)	1899 (32.48)	183 (3.13)	1326 (22.68)	102 (1.74)	1673 (28.61)	620 (10.60)	5847
Europe	1530 (1.60)	9444 (9.89)	2797 (2.93)	46,765 (48.99)	1968 (2.06)	28,580 (29.94)	4372 (4.58)	95,456
North America	949 (0.93)	9652 (9.46)	2809 (2.75)	27,176 (26.63)	2073 (2.03)	56,358 (55.23)	3042 (2.98)	102,069
Oceania	14 (0.33)	283 (6.72)	93 (2.21)	1645 (39.04)	36 (0.85)	1812 (43.00)	331 (7.85)	4214
Total	2537 (1.22)	21,278 (10.25)	5882 (2.83)	76,912 (37.05)	4179 (2.01)	88,433 (42.60)	8365 (4.03)	207,586

^aPercentages are based on row totals and may not sum to 100 because of rounding.

Design A sample of initial results submissions were first reviewed by ClinicalTrials.gov staff per standard review procedures, then a second reviewer examined a convenience subsample to assess agreement and categorize major issues using categories derived from the results review criteria. Major issues were only counted once per submission.

Results Among 358 initial results submissions in the 4 weeks between July 19, 2015 and August 15, 2015, 240 had major issues. In a convenience subsample of 215 submissions (114 [53%] nonindustry and 101 [47%] industry) we identified 471 occurrences of 37 unique major issue categories with a mean (SD) of 2.2 (1.3) unique major issues overall (1.9 [1.2] industry and 2.5 [1.4] non-industry). The top 12 major issue categories accounted for 398 (85%) occurrences of all major issues (**Table 55**). The top 5 unique major issues (occurring in ≥20%) in submissions had an invalid and/or inconsistent unit of measure (86 [40%]), insufficient information about a scale (55 [26%]), internal inconsistency (52 [24%]), written results or conclusions (47 [22%]), and unclear baseline or outcome measure (44 [20%]).

Conclusions Most major issues identified in a convenience sample of results submissions at ClinicalTrials.gov could be described with only 12 categories. Limitations of this analysis include the use of a convenience sample and assessment of major issues by 2 people sequentially rather than independently. Further research is needed to confirm the generalizability of these findings, with an aim of improving the validation process, developing targeted support materials, and improving results reporting on the platform.

^aClinicalTrials.gov National Library of Medicine, National Institutes of Health, 8600 Rockville Pike, Bethesda, MD, USA, heather.dobbins@nih.gov

Table 55. Categories of Major Issues Occurring in Results Submissions^a

Major Issue	215 Results Submissions, No. (%)
Invalid and/or inconsistent unit of measure	86 (40)
Insufficient information about a scale used for assessment	55 (26)
Internal inconsistency (inconsistency between information in different parts of the record)	52 (24)
Written results or conclusions in free text that are not in tabular format	47 (22)
Unclear baseline or outcome measure	44 (20)
Incorrect Measure Type	23 (11)
Zero participants at risk for adverse events without explanation	19 (9)
Data with multiple units of measure	19 (9)
Multiple time points without an explanation	15 (7)
Results not reported for each arm separately	14 (7)
Nonmeaningful values included as "placeholder" data	12 (6)
Adverse events at risk population inconsistent with other information in record	12 (6)

^aA total of 471 occurrences of 37 unique major issues.

Conflict of Interest Disclosures: All authors report working for ClinicalTrials.gov. Dr Zarin is a member of the Peer Review Congress Advisory Board but was not involved in the review or decision of this abstract.

Funding/Support: Supported by the Intramural Research Program of the National Library of Medicine, National Institutes of Health.

Role of the Funder: Four authors are full-time employees (HDD, TT, RJW, DZ), and 1 author is a contractor (CC) of the National Library of Medicine. The National Library of Medicine has approved this submission.

MONDAY

Transparency in Cross-National Research: Quality of Reporting

Elena Damian,¹ Bart Meuleman,¹ Wim van Oorschot¹

Objective This study investigates the degree of transparency in cross-national research. It provides insight into what information is most likely to be left out in empirical studies and to what extent researchers provide sufficient details to assess the quality of their studies and/or to possibly replicate them.

Design The data set is composed of 305 comparative studies that were published between 1986 and 2016 in 1 of 29 sociology, political science, and cross-cultural psychology journals. First, we selected all journals from these fields that accept manuscripts on a broader variety of topics and publish comparative research. Second, we selected all articles that (1) use data from at least 1 of 7 international surveys that offer free data access (ie, Afrobarometer, Eurobarometer, European Social Survey, European Values Study, International Social Survey Program, World Values Survey, and Latinobarometer); (2) include in their analyses 5 or more countries; (3) use any type of comparative analysis; and (4) do not have a purely methodological aim. This selection resulted in 1007 studies from which we drew a random sample of 305 articles. Third, we created a questionnaire and coded for each article what information regarding the empirical analysis is reported (eg, sampling design, description and measurement of the variables, and information about the data used for contextual variables).

Results We found that most studies include basic information about the empirical analysis: a description of the population sample (81%, 246) and the dependent variables (97%, 297), and the contextual variables (95%; n = 191 of 202 articles with variables from external sources); the exact questions used to measure the dependent variables (70%, n = 212 of 305); a list of the countries included in the study (89%; n = 271 of 305); or the final study sample size (82%; n = 249 of 305). However, less than half of the articles provide crucial information needed to assess the quality of the study, ie, information about the sample design (39%; n = 118 of 305), survey mode (16%; n = 50 of 305), response rate (9%; n = 28 of 305), use of weights (29%; n = 87 of 305), number of missing values (10%, n = 30 of 305) and their treatment (31%;

n = 94 of 305), precise references to the data sources used to create the contextual variables (42% n = 85 of 202), or dataset version (18%; n = 55 of 305). In addition, of all 305 articles analyzed, only 2 articles provided full and accessible replication materials.

Conclusions These preliminary results reveal that most cross-national studies published in sociology, political science, and cross-cultural psychology journals omit essential information needed to assess their quality.

¹Centre for Sociological Research, University of Leuven, Leuven, Belgium, elena.damian@kuleuven.be

Conflict of Interest Disclosures: None reported.

Funding/Support: No external funding outside of the employer of the 3 authors (University of Leuven).

Role of the Funder/Sponsor: The sponsor of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

TUESDAY

Presence of a Unique Trial Identifier in the Abstracts of Industry-Sponsored Manuscripts

LaVerne A. Mooney,¹ Joseph F. Michalski,² Lorna Fay¹

Objective Since 2008, the International Committee of Medical Journal Editors (ICMJE) has recommended that journals include a unique trial identifier (eg, the ClinicalTrials.gov ID number [NCT]) in the abstract of clinical trial manuscripts. Adherence to this recommendation should result in automatic linkage between the PubMed abstract and the trial record on ClinicalTrials.gov, and inclusion of the manuscript citation in the trial record. A 2013 study reported that unique trial identifiers are frequently missing from published manuscript abstracts. Our objective was to assess the implementation of the ICMJE’s recommendation in manuscripts reporting results of Pfizer-sponsored trials and determine whether the relevant citation was present on ClinicalTrials.gov.

Design Using a Pfizer publication database, we identified manuscripts reporting primary outcomes for Pfizer clinical trials published from 2013 to 2015, and we obtained

corresponding NCT numbers from ClinicalTrials.gov. We excluded manuscripts of trials that were not indexed on PubMed, that were not registered on ClinicalTrials.gov, or that reported non-interventional studies. For each clinical trial, we recorded the presence or absence of the NCT number in the manuscript PDF and its location within the manuscript, the presence or absence of the NCT number in the PubMed abstract, and the presence or absence of a manuscript citation on the trial record and if it had been automatically indexed or added by the sponsor. We report percentages overall and by year.

Results A total of 276 manuscripts (305 studies) were published in 140 unique journals. Half of the PubMed abstracts (143 of 276 [51.8%]) included the NCT number, and the citation was present in 142 of 276 (51.4%) of the ClinicalTrials.gov records (**Table 56**). Auto-indexing accounted for 140 of 142 (99.0%) of the citations on trial records; only 2 resulted from manual addition by the sponsor to the study records when the NCT number was not present in the abstracts. Errors or failures of the auto-indexing process appeared to occur for 3 abstracts published in 2013.

Conclusions The NCT number was included in most Pfizer clinical trial manuscripts (232 of 276 [84.1%]), but auto-indexing and bidirectional linkage (from PubMed and from ClinicalTrials.gov) only occurred when the NCT number was located in the abstract per ICMJE guidelines. An additional 89 manuscript citations could have been visible to the public on ClinicalTrials.gov if the NCT number was correctly positioned in the abstract. This study demonstrates the importance of the location of the trial identifier within the manuscript.

¹Pfizer Medical, New York, NY, USA, laverne.mooney@pfizer.com; Mailman School of Public Health, Health Policy and Management, Columbia University, New York, NY, USA

Conflict of Interest Disclosures: L. Mooney and L. Fay are employees of Pfizer Inc and hold Pfizer stock. J. Michalski is a graduate student who assisted with the research, was paid by Pfizer, and holds Pfizer stock. No other disclosures were reported.

Funding/Support: Pfizer is the employer of two of the authors and the analysis was carried out by them at Pfizer.

Table 56. Data on Location of NCT Numbers in PubMed Abstracts and Manuscripts and Linked Citations on ClinicalTrials.gov

Year	Manuscripts, No. ^a	NCT Number, No. (%)		NCT Number Not in PubMed Abstract but Present in Manuscript PDF, No. (%)	Manuscript Citation in ClinicalTrials.gov Record, No. (%) ^c	Trial Records and Citations Auto-Linked by NCT Number, No. (%)	
		In Manuscript PDF	In PubMed Abstract ^b			NCT Number in Abstract ^d	NCT Number Not in Abstract
2013	99	78 (78.8)	53 (53.5)	25 (32.0)	50 (50.5)	50 (50.5)	0
2014	105	94 (89.5)	60 (57.1)	34 (36.2)	60 (57.1)	60 (57.1)	0
2015	72	60 (83.3)	30 (41.7)	30 (50.0)	32 (44.4)	30 (41.7)	0
All	276	232 (84.1)	143 (51.8)	89 (38.4)	142 (51.4)	140 (50.7)	0

Abbreviation: NCT, The ClinicalTrials.gov ID number.
^aThe results of 305 trials were published in 276 manuscripts (23 manuscripts reported >1 trial) and included corresponding NCT numbers.
^bAll NCT numbers for the trials reported in a manuscript had to be included in the abstract to be considered present.
^cTwo citations were added manually by the sponsor of 2 trials to the ClinicalTrials.gov database, not auto-indexed by NCT number.
^dFor manuscripts reporting multiple trials, the citation had to be present on each trial record via auto-indexing to be included.

Assessment of the Quality and Transparency of Research Reporting Endorsement by Brazilian Health Science Journals

Tais F. Galvao,¹ Monica C. Roa,² Leila Posenato Garcia,³ Marcus T. Silva⁴

Objective To assess the quality and transparency of research reporting endorsement by Brazilian health science journals.

Design We assessed Brazilian MEDLINE-indexed journals by applying the 10 questions used by Moher et al (2016) to investigate journals' efforts to increase value and reduce waste in biomedical research. Journals were selected through searching the Abridged Index Medicus (via NLM Catalog) using the strategy "currentlyindexed[All] AND (Brazil* OR Brasil*)." Eligible health science journals were selected from the retrieved titles. We reviewed journals' instructions to authors, editorials, and articles published between 2014 and 2016 to answer to the 10 questions. We also checked the journals' citation metrics (Impact Factor at InCites Journal Citation Reports, SCImago Journal Rank, and journals' H-index), indexing bibliographic databases, publication language, costs, and periodicity. Data were extracted, inserted into a previously designed spreadsheet, and analyzed by descriptive statistics. We used linear regression to test the association between bibliometric measures and the 10-question score for reducing research waste.

Results The search retrieved 60 journal titles, and 10 were excluded: 5 not Brazilian, 4 related to basic sciences, and 1 duplicated. Of the 50 journals investigated, 27 had an Impact Factor (median, 0.937; interquartile range, 0.730-1.194), 48 had a SCImago Journal Rank (median, 0.356; interquartile range, 0.268-0.528) and 48 informed their H-index (median, 22; interquartile range, 12-35). Half (25) publish bimonthly. No submission or publication fee is required by 33 journals. Other indexing databases were, in order of frequency: LILACS, Scopus, and SciELO. No Brazilian journal met all 10 questions on efforts to increase value and reduce waste (Table 57). The median of items fulfilled was 5 (maximum, 8 [2 journals]; minimum, 1 [1 journal]). All except 1 of the journals provide support for good reporting infrastructure, 46 are open access or have policies for it, 39 recommend authors visit the International Committee of Medical Journal Editors website, 32 mention mandatory registration of clinical trials or systematic reviews, 14 mention the use of reporting guidelines, and 3 mention the Enhancing the Quality and Transparency of Health Research (EQUATOR) Network. None mention systematic reviews as part of main study results. There was a significant negative association between journal's SCImago Journal Rank and scores for reducing research waste ($\beta = -2.73$; $P = .04$). There were no significant associations between the journal scores and the other bibliometric measures.

Conclusions Brazilian health science journals partially encourage quality and transparency in research reporting. Endorsement of the EQUATOR Network and reporting

guidelines in the instructions to authors are simple efforts that would increase the value of published research.

¹Faculty of Pharmaceutical Sciences, State University of Campinas, Campinas, Brazil, taisgalvao@gmail.com; ²Faculty of Medicine, Clinical Research Institute, National University of Colombia, Bogotá, Colombia; ³Institute of Applied Economic Research, Brasília, DF, Brazil; ⁴Faculty of Medicine, Federal University of Amazonas, Manaus, Brazil

Conflict of Interest Disclosures: None reported.

Funding/Support: None reported.

Reporting Statistical Inference in Psychiatry 1975-2015: An Analysis of Abstracts in Major Psychiatric Journals

Christopher Baethge,^{1,2} Markus Deckert,³ Andreas Stang^{3,4}

Objective We analyzed changes in the reporting of statistical inference in the abstracts of psychiatric journals, specifically comparing changes over time in the reporting of *P* value thresholds (eg, $P \leq .05$ or use of the term *significant*) vs precise *P* values and the use of confidence intervals.

Design Using an SAS-based algorithm for PubMed, we searched 63,928 abstracts published between 1975 and 2015 in 15 leading psychiatric journals (defined by 2015 impact factors). For validation, we checked 950 randomly selected abstracts.

Results The mean (SD) length of abstracts increased from 716 (317) characters in 1975-1979 to 1928 (1113) characters in 2010-2015, and the use of numbers increased from a median

Table 57. Brazilian Health Science Journal Efforts to Increase Value and Reduce Waste in Biomedical Research (N=50)

Questions	No. (%)
Does the journal's instructions to authors explicitly mention reporting guidelines?	17 (34)
Does the journal's instructions to authors explicitly mention the Enhancing the Quality and Transparency of Health Research Network?	3 (6)
Does the journal's instructions to authors explicitly mention clinical trial, systematic review, or other registration?	31 (62)
Does the journal's instructions to authors mention use of systematic reviews as part of reporting main study results?	0
Does the journal's instructions to authors recommend authors to go to the website of the International Committee of Medical Journal Editors for guidance?	38 (76)
Does the journal support publishing research on research, such as methods and reporting section?	22 (44)
Has the journal published editorials about the Series, other pieces on waste, duplication, reporting guidelines, registration, or other topics related to increasing research value?	27 (54)
Does the journal provide support for good reporting infrastructure?	49 (98)
Does the journal mention anything or have policies about open access?	47 (94)
Does the journal have a policy on public access to data from completed research?	4 (8)

of 2 per abstract (10th to 90th percentile range, 0-10 per abstract) to 14 per abstract (10th to 90th percentile range, 0-72 per abstract). An estimated 3.6% (95% CI, 2.5%-5.2%) of abstracts contained nonstatistical significance terminology only. About 45% of all abstracts reported some form of statistical inference, increasing from 26% in 1975-1979 to 52% in 2010-2015. In those abstracts, statistical inference based on *P* value thresholds was dominant (**Figure 12**), decreasing from 99% (1975-1979) to 66% (2010-2015); but with more articles being published, the absolute numbers that reported *P* value thresholds increased from 1095 in 1975-1979 to 3806 recently. While reporting precise *P* values did not appear at all 40 years ago and remained rare until recently (6% of abstracts in 2010-2015), combining precise and threshold *P* values is now more common (1% of abstract in 1975-1979 vs 20% of abstracts in 2010-2015). From 2010 to 2015, 22% of abstracts included confidence intervals and 7% displayed confidence intervals only; from 1975 to 1979, no abstracts included confidence intervals. All results varied widely across journals.

Conclusions In the abstracts of psychiatric articles, we detected a shift from reporting *P* value thresholds only (eg, $P \leq .05$) to presenting precise *P* values and confidence intervals, although the use of *P* value thresholds and the absence of confidence intervals remain common. Analyzing full-text articles might provide more complete information about these trends, but the findings suggest that the decades-long debate on “estimation over testing” has not gained much ground among researchers in psychiatry and has not led to a substantial replacement of *P* values by confidence intervals.

¹Department of Psychiatry and Psychotherapy, University of Cologne Medical School, Cologne, Germany; ²*Deutsches Ärzteblatt* and *Deutsches Ärzteblatt International*, Editorial Offices, Cologne, Germany, baethge@aerzteblatt.de; ³Center of Clinical Epidemiology,

c/o Institute of Medical Informatics, Biometry, and Epidemiology, University Hospital of Essen, Essen, Germany; ⁴Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA

Conflict of Interest Disclosures: None reported.

TUESDAY

A Scale for the Assessment of Non-systematic Review Articles (SANRA)

Christopher Baethge,^{1,2} Sandra Goldbeck-Wood,^{3,4} Stephan Mertens¹

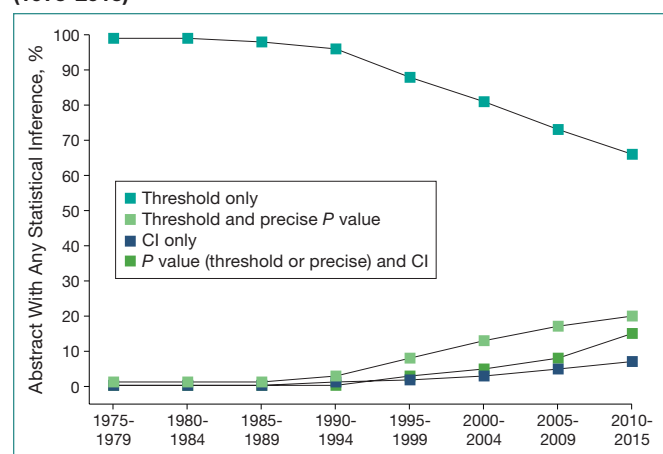
Objective To revise the Scale for the Assessment of Non-systematic Review Articles (SANRA), an instrument developed to help editors, reviewers, and researchers assess the quality of non-systematic review articles, and to test it in a larger number of manuscripts.

Design A team of 3 journal editors modified items in an earlier SANRA version based on face validity, item-total correlations, and reliability scores from previous tests, and deleted an item addressing a manuscript’s writing and accessibility because ratings differed considerably. The revised scale comprises 6 items scaled from 0 (low standard) to 2 (high standard) related to (1) justification of the review’s importance, (2) aims of the review, (3) literature search description, (4) adequacy of referencing, (5) presentation of levels of evidence, and (6) presentation of data central to the article’s argument. For all items we developed recommendations and examples to guide users filling out the instrument. The revised scale was tested by the same editors, blinded to each other’s ratings, in a group of 30 consecutive non-systematic review manuscripts submitted to *Deutsches Ärzteblatt*, a general medical journal, in 2015.

Results The mean (SD) sum score across the 30 manuscripts was 6.0 (2.6) [range, 1-12]. Corrected item-total correlations ranged from 0.33 (item 3) to 0.58 (item 6). Cronbach $\alpha = .68$. The intraclass correlation coefficient (average measure) was 0.77 (95% CI, 0.57-0.88). Raters often disagreed on items 1 and 4. Raters confirmed that completing the scale in approximately 5 minutes is feasible in everyday editorial work and that it is easier to understand than the earlier version.

Conclusions A revised 6-item version of SANRA, a rating scale for the assessment of non-systematic reviews, demonstrated interrater reliability, homogeneity of items, and internal consistency sufficient for a scale of 6 items. In comparison with earlier versions of the scale, the current version is shorter, is based on appropriate field tests, and is easier to use. Further testing of the scale’s validity (eg, expert ratings of manuscripts, citations, reviewer recommendations) is desirable, as is rater training based on recommendations and examples provided with the scale. The scale is intended to complement rather than replace journal-specific evaluation of manuscripts (eg, pertaining to audience, originality or difficulty) and may contribute to improving the standard of reporting of non-systematic reviews.

Figure 12. Statistical Inference in 15 Psychiatric Journals (1975-2015)



The journals included were *World Psychiatry*, *JAMA Psychiatry*, the *American Journal of Psychiatry*, *Molecular Psychiatry*, *Biological Psychiatry*, the *Schizophrenia Bulletin*, *Psychotherapy and Psychosomatics*, the *Journal of the American Academy of Child and Adolescent Psychiatry*, the *British Journal of Psychiatry*, the *Journal of Child Psychology and Psychiatry*, the *Journal of Neurology and Neurosurgery*, *Neuropsychopharmacology*, *Acta Psychiatrica Scandinavica*, *Lancet Psychiatry*, and the *Journal of Psychiatry and Neuroscience*.

¹Deutsches Ärzteblatt and Deutsches Ärzteblatt International, Cologne, Germany, baethge@aerzteblatt.de; ²Department of Psychiatry and Psychotherapy, University of Cologne Medical School, Cologne, Germany; ³Department of Obstetrics and Gynecology, Addenbrooke's Hospital, Cambridge University, Cambridge, UK; ⁴Journal of Family Planning and Reproductive Health Care, London, UK

Conflict of Interest Disclosures: None reported.

TUESDAY

Completeness of Reporting in Indian Qualitative Public Health Research: A Systematic Review of 20 Years of Literature

Myron A. Godinho,¹ Nachiket Gudi,² Maja Milkowska,³ Shruti Murthy,⁴ Ajay Bailey,⁵ N. Sreekumaran Nair⁵

Objective To systematically review the completeness of reporting in published, qualitative public health research (QPHR) studies in India.

Design An electronic search was conducted in PubMed, Ovid MEDLINE, Web of Science, SCOPUS, and GoPubMed to identify English-language primary QPHR studies from India, published between January 1, 1997, and June 30, 2016. Study selection was based on title, abstract, and full-text reviews. To assess included articles, the following modifications were made to the Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist: item 9 was subdivided into 3 subcategories (study design, methodological orientation, data collection method); item 33 was added (study limitations); and all original items were retained. All included articles were independently assessed for compliance with the COREQ checklist by 2 groups of 2 reviewers, and each item was noted as either reported or unreported. The summary included the number of COREQ items reported by each study and the

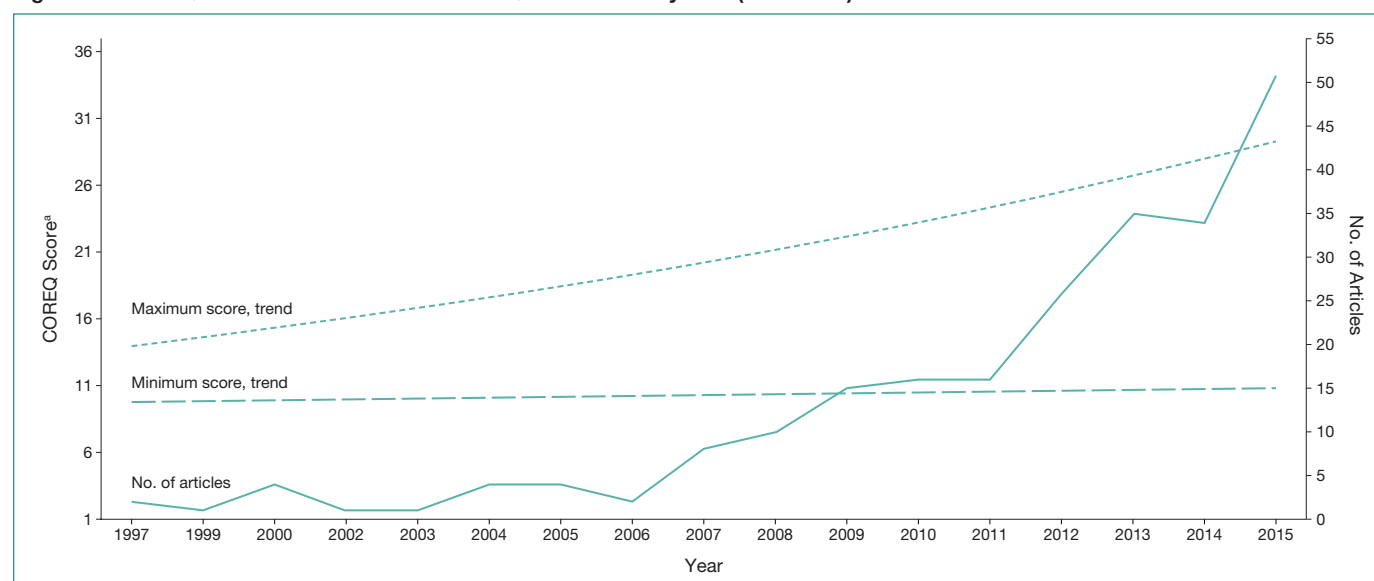
number of studies that reported each COREQ item. Descriptive statistics for each year and for pre- and post-COREQ time periods were reported.

Results Following assessment of 893 citations, 246 articles were included. Trends demonstrated an increasing number of Indian QPHR studies being published annually and an increasing maximum number of COREQ items reported over the last 20 years (**Figure 13**). However, there was no increase in the minimum number of reported COREQ items over this period. Only 2 COREQ items (study design, and data collection method) were reported in all 246 studies. More than half of the studies reported 16 to 21 of the overall 37 items. The least-reported items (reported in <10% of studies) were mention of repeat interviews (1.6%), specification of reasons for nonparticipation (2.8%), review of interview transcripts by participants (4.5%), description of coding tree (6.1%), relationship establishment prior to study commencement (8.1%), and participant checking (8.9%). Despite a net percentage increase in reporting in all 3 domains after the introduction of COREQ (in 2007), domain 1 (research team and reflexivity) was least-frequently reported, with the least percentage improvement.

Conclusions Most studies reported approximately half of the recommended items. Despite improving trends, the reporting of QPHR in India is far from complete. The interviewers, context, and analysis can all affect study conclusions, and hence should be reported. Authors, journal editors, and other involved individuals should collaborate to ensure adherence to established reporting guidelines, with attention focused on the least-frequently reported areas.

¹Public Health Evidence South Asia (PHESA), Manipal University, Manipal, India, myrongodinho@gmail.com; ²School of Allied Health Sciences, Manipal University, Manipal, India; ³Maastricht University, Maastricht, the Netherlands; ⁴Public Health Evidence

Figure 13. COREQ Scores and Number of Indian QPHR Articles by Year (1997-2015)



^aThe Consolidated Criteria for Reporting Qualitative Research (COREQ) score assigned to an article directly corresponds to the number of items reported in that article.

Minimum and maximum score trends indicate linear time trends in the lowest scored article and the highest scored article in a given year. QPHR indicates qualitative public health research.

South Asia (PHESA), Manipal University, Manipal, India;
⁵Population Research Centre, Faculty of Spatial Sciences, University of Groningen, Groningen, the Netherlands; ⁶Public Health Evidence South Asia (PHESA), Department of Statistics, Manipal University, Manipal, India

Conflict of Interest Disclosures: None reported.

Funding/Support: No funding was provided or obtained for any part of this study.

Acknowledgments: We would like to acknowledge Public Health Evidence South Asia (PHESA), and Manipal University for providing the infrastructural support necessary to conduct this research.

MONDAY

Reporting of Sex and Gender in Clinical Trial Protocols and Published Results

Thiyagu Rajakannan,¹ Kevin M. Fain,¹ Rebecca Williams,¹ Tony Tse,¹ Deborah A. Zarín¹

Objective Biomedical research funders and journals have increasingly focused on the importance of assessing and reporting the effect of sex (biological factors) and gender (sociocultural factors) on health outcomes in clinical studies. Prior literature reviews have indicated that sex or gender are frequently not assessed or reported in published clinical studies. These studies did not assess research designs in protocols. The object of this study was to assess publicly available clinical study protocols and corresponding published studies to analyze how “sex” and “gender” information was incorporated in study design and reported.

Design We identified a convenience sample of 80 articles from *New England Journal of Medicine (NEJM)* and *JAMA* published in 2014–2015 for which full protocols were available online. We then searched for and assessed the use of the terms “sex” and “gender” in the entire protocol and corresponding article.

Results We found that, first, the terms “sex” and “gender” were not defined in any of the protocols or articles. Second, of the 80 clinical trials analyzed, 32 (40%) used both terms interchangeably in the protocol; 28 of these used “sex” only, and 4 used neither term in the corresponding article (**Table 58**). No article used the terms interchangeably. Finally, the term “gender” only was used in 23 (29%) protocols, but only 1 article used the term “gender.” Our data indicate imprecision in the use of the terms “sex” and “gender” in study protocols, suggesting a lack of appreciation among researchers of these distinct concepts. Articles generally used only “sex,” implying that journals enforce the use of specific and consistent terminology when reported. Of the 80 included studies, 14 (18%) articles did not use the terms “gender” or “sex” and were not sex-specific studies, of which 10 of these 14 (13% overall) used terms such as “men” or “women” but were unclear whether gender or sex was meant. We note the generalizability of these findings may be limited. These journals were chosen because they systematically included protocols for a large sample of clinical trials. The *JAMA* instructions for authors specifically

Table 58. Sex/Gender Reporting Anywhere in 80 Selected Trial Protocols and Articles (40 in NEJM and 40 in JAMA)^a

Terms Used	No. (%)
Clinical Trial Protocols	
Protocols that use the term “sex” only	11 (14)
Protocols that use term “gender” only	23 (29)
Protocols that use both “sex” and “gender” interchangeably	32 (40)
Protocols that use neither term	14 (17)
Sex-specific studies (such as breast cancer)	4 (5)
Not sex-specific studies (other phrase used)	7 (9)
Not sex-specific studies (no other phrase used)	3 (4)
Articles	
Articles that use the term “sex” only	59 (74)
Articles that use “gender” only	1 (1)
Articles that use both “sex” and “gender” interchangeably	0
Articles that use neither term	20 (25)
Sex-specific studies (such as breast cancer)	6 (7)
Not sex-specific studies (other phrase used)	10 (12)
Not sex-specific studies (no other phrase used)	4 (5)

^aPercentages may not equal 100 owing to rounding.

addressed sex/gender reporting, although the *NEJM* author instructions did not explicitly address this reporting. Also, we did not assess how the constructs were used in the research.

Conclusions Our study supports the need for continuing efforts to standardize the concepts of “sex” and “gender” and ensure their appropriate use in biomedical research.

¹ClinicalTrials.gov, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA, kevin.fain@nih.gov

Conflict of Interest Disclosures: All authors report working on ClinicalTrials.gov. Dr Zarín is a member of the Peer Review Congress Advisory Board but was not involved in the review or decision of this abstract.

Funding Support: Supported by the Intramural Research Program of the National Library of Medicine, National Institutes of Health.

Role of the Funder: All authors are full-time employees of the National Library of Medicine. The National Library of Medicine has approved this submission.

TUESDAY

Adherence to Consolidated Standards of Reporting Trials (CONSORT) Guideline Items in Randomized Trials of Physical Activity Published in 5 Sports Medicine Journals

Daniel Umpierre,^{1,2,3} Lucas Helal,¹ Patrícia Martins Bock,^{1,4} Lucas Porto Santos¹

Objective Physical activity trials are being published more often, but reports may not include randomized clinical trial reporting guideline items. We aimed to describe the completeness of guideline-recommended reporting in

randomized clinical trials of supervised or unsupervised physical activity interventions.

Design We conducted MEDLINE searches to identify RCTs published in 5 leading journals in sports and exercise medicine (*British Journal of Sports Medicine*, *Medicine and Science in Sports and Exercise*, *Scandinavian Journal of Medicine and Science in Sports*, *Journal of Applied Physiology*, and *International Journal of Sports Medicine*) between January 2006 and December 2016. Eligible studies were primary reports having at least 1 intervention arm comprising structured physical activity (eg, tailored exercise programs) or unstructured physical activity programs (eg, general exercise recommendations). We excluded studies using a physical activity intervention arm as a comparator for a nonphysical activity active intervention. For each publication, 2 assessors independently reviewed full texts and abstracted data based on 9 items selected from Consolidated Standards of Reporting Trials (CONSORT) guidelines. The 9 items were chosen based on their applicability to physical activity trials and their relevance to readers (eg, title identification) and researchers (eg, numbers analyzed). Our review is ongoing; the findings reported here derive from a 25% random sample of retrieved articles for each journal.

Results Of 86 randomized clinical trials, 40 (47%) and 46 (53%) were published in the periods ranging from 2006 to 2016, respectively, 14 (16%) were identified as a randomized trial in the title; 65 (76%) provided details on trial design and allocation ratio; 50 (58%) mentioned specific hypotheses; 22 (26%) stated their primary outcome; 16 (19%) described sample size calculation; 23 (27%) mentioned examiners blinding; 14 (16%) explicitly indicated the number of participants analyzed; 20 (23%) mentioned trial registry; and only 6 (7%) cited an accessible full trial protocol. One trial reported all 9 items, 2 reported 8, and 7 did not report any item.

Conclusions A low to modest proportion of trials published in 5 leading exercise science journals reported important CONSORT guideline items. Because careful trial description can enhance completeness, transparency, and reproducibility in exercise sciences, journal endorsement of standard reporting guidelines or checklists might be desirable.

¹Exercise Pathophysiology Research Laboratory, Federal University of Rio Grande do Sul, Porto Alegre, Brazil, daniel.umpierre@gmail.com; ² National Institute of Science and Technology for Health Technology Assessment (IATS)-CNPq, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil; ³School of Nursing, Federal University of Rio Grande do Sul, Porto Alegre, Brazil; ⁴Faculdades Integradas de Taquara, Rio Grande do Sul, Brazil

Conflict of Interest Disclosures: None reported.

Funding/Support: Daniel Umpierre receives research productivity grants from the National Council on Technology and Scientific Development (CNPq, Brazil). Patrícia Martins Bock receives postdoctoral scholarship from the Coordination for the Improvement of Higher Level Personnel (Capes). The other authors report no disclosures.

Role of the Funder/Sponsor: The sponsors of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Additional Information: Datasets are publicly available at <https://github.com/dumpierre/exsc-reporting.git>

TUESDAY

Identification of Ethics Committees Based on Authors' Disclosures: A Cross-Sectional Study of Articles Published in the *European Journal of Anaesthesiology* and a Survey of Ethics Committees

Davide Zoccatelli,¹ Martin R. Tramèr,^{1,2} Nadia Elia^{1,3}

Objective In July 2010, the *European Journal of Anaesthesiology* (*EJA*) began to progressively implement requests to authors to specify 5 items related to ethical approval (EA) in their manuscripts: name and address of the responsible ethics committee (EC), name of the chairperson, protocol number, and date of approval. We sought to assess if provision of these details facilitated identification of, contact with, and confirmation of approval by ECs.

Methods We identified all articles published in *EJA* in 2011 that required EA according to the Swiss Federal Act on Research Involving Human Beings, focusing on the year 2011 because the reporting of EA was still heterogeneous. From each of the included studies we extracted which and how many of the required items were reported, and attempted to identify the EC based on the reported information in the EA declaration. We contacted each identified EC to seek confirmation of their role in the EA of the respective studies. We compared proportions of ECs identified and number of ECs confirming their role for studies reporting 5 items vs those reporting 4 or fewer using a χ^2 test (α , 0.05; bilateral).

Results Of 193 articles published in 2011, 76 required and 74 (97%) declared EA. The name and address of an EC were mentioned in 63 (85%) EA declarations, protocol number in 51 (69%), date of approval in 48 (65%), and name of a chairperson in 45 (59%). All 5 items were reported in 34 (47%) articles, and 4 or fewer items were reported in 40 (53%) articles. We were able to identify and contact 44 (59%) ECs; 36 (48%) eventually responded, of which 24 (32%) confirmed approval, 10 (14%) were unable to confirm approval for a variety of reasons, and 2 (3%) refused to confirm approval for legal reasons. Reporting all 5 items, compared with 4 or fewer, increased the chance that an EC would be identified ($P = .02$), and would confirm its approval ($P = .05$).

Conclusion The reporting of 5 items related to EA facilitates the identification of the competent EC, and increases the likelihood that the EC would confirm having given EA. However, identification of, and successful contact with EC remains difficult. Future research should identify which information could allow easier and successful identification and contact with ECs.

¹Division of Anesthesiology, Department Anesthesiology, Pharmacology and Intensive Care Medicine, Geneva University Hospitals, Geneva, Switzerland, martin.tramer@hcuge.ch; ²Faculty of Medicine, University of Geneva, Geneva, Switzerland; ³Institute of Global Health, University of Geneva, Geneva, Switzerland

Conflict of Interest Disclosures: Dr Tramèr is a member of the Peer Review Congress Advisory Board but was not involved in the review or decision for this abstract. No other conflicts are reported.

Quality of the Literature

MONDAY

The Role of Supplementary Material in Journal Articles: Surveys of Authors, Reviewers, and Readers

Amy Price,^{1,2} Sara Schroter,¹ Mike Clarke,^{3,4,5} Helen McAneney⁵

Objective Many journals allow or require authors to submit supplementary material for consideration when their manuscript is going through the editorial process and for possible publication with the article. We explore the value and role of supplementary material in journal articles from the perspective of authors, peer reviewers, and readers.

Design Among authors and peer reviewers of research submissions to 17 BMJ Publishing Group journals, we randomly allocated two-thirds of each group to receive an author and reviewer survey, respectively, and the remaining third of each group to receive a reader survey. In November 2016, participants completed an online survey from the perspective of their allocated role to provide information about their use of specific types of supplementary material (study protocol, data collection or extraction forms, data tables and figures, completed reporting guideline checklists and flow diagrams, interview transcripts, raw study data). Survey questions asked about who each each portion of the material is most useful to; the expected use of materials by authors, reviewers and readers; the preferred option for accessing supplementary material; and if and where supplementary material should be published.

Results Among 20,340 surveyed authors and peer reviewers, we received 2872 (14%) responses (819 [12%] from authors, 1142 [17%] from peer reviewers, and 911 [14%] from authors and reviewers responding as readers). Most authors (711 of 819 [87%]) reported submitting at least 1 type of supplementary material with their most recent manuscript, 95% (1086 of 1142) of reviewers reported seeing supplementary material in an article they reviewed at least sometimes (ie, not *never* or *almost never*), and 79% (724 of 911) of readers reported that the article they read most recently included supplementary material. Additional data tables were the most common supplementary material type submitted or seen (authors, 74%; reviewers, 89%; readers, 67%). A majority in each role indicated additional data tables were most useful to readers (58%-72%), while fewer indicated they were most useful to peer reviewers (19%-34%) and journal editors (3%-4%) (**Table 59**). Patterns of opinion were

Table 59. Author, Reviewer, and Reader Perspectives on the Value of Additional Tables of Data and Completed Checklists for Reporting Guidelines by Group^a

Group	No./Total No. Most Useful (%)		
	To Journal Editors	To Peer Reviewers	To Readers
Additional tables of data			
Authors	29/819 (4)	187/819 (23)	564/819 (69)
Reviewers	32/1142 (3)	384/1142 (34)	662/1142 (58)
Readers	25/911 (3)	172/911 (19)	659/911 (72)
Completed checklists for reporting guidelines			
Authors	365/819 (45)	291/819 (36)	96/819 (12)
Reviewers	453/1142 (40)	414/1142 (36)	186/1142 (16)
Readers	340/911 (37)	394/911 (43)	117/911 (13)

^aPercentages do not sum to 100% because of missing data.

opposite for reporting guideline checklists. All 3 groups favored publication of additional data tables and figures on the journal’s website (80%-83%), with less than 4% of each group reporting that availability was not needed. Less than one-fourth of respondents in each group said that raw study data should be available on the journal’s website (16%-23%), and 24% to 33% said that these materials should not be made available.

Conclusions Authors, peer reviewers, and readers favor access to supplementary tables and figures over completed reporting checklists or raw data. These findings may help journals to consider the roles, resource costs, and strategic placement of supplementary materials for optimal usage.

¹The BMJ, London, UK, dr.amyprice@gmail.com; ²Department of Continuing Education, University of Oxford, Oxford, UK; ³Northern Ireland Methodology Hub, Belfast, UK; ⁴Northern Ireland Clinical Trials Unit, Belfast, UK; ⁵Centre for Public Health, Queen’s University Belfast, Belfast, UK

Conflict of Interest Disclosures: Amy Price is a research fellow at and Sara Schroter is a full time employee of The BMJ. Mike Clarke reports involvement in many clinical trials and systematic reviews and has prepared and used supplementary material widely. He seeks funding for these trials and reviews, as well as for research into methodology, including dissemination and accessibility. No other disclosures were reported.

Acknowledgments: We would like to acknowledge the 45 volunteers who piloted this research and shared valuable feedback to make the questions clear. The volunteers were community members, physicians, researchers, patients, and teachers. We thank all the researchers who completed the survey and especially those who shared comments. Their perspectives have increased our understanding.

MONDAY

Editorial Evaluation, Peer Review, and Publication of Research Papers With and Without Online-Only Supplements: Quality vs Superior Tonnage

Annette Flanagan,¹ Stacy Christiansen,¹ Chris Borden,¹ Demetrios Kyriacou,¹ Caroline Sietmann,¹ Elaine Williams,¹ Larry Bryant,¹ Jamie Reickel,² Allison Newton²

Objective Online-only supplements are a useful option for publication of content or data that may not fit in the space allotted to a scientific article, but the value of these supplements is not known. We conducted this study to determine if research papers with supplements have a higher likelihood of submission, peer review, acceptance, and online usage after publication than those without supplements.

Design Cohort study that assessed the numbers and types of supplements submitted with research manuscripts to the 3 medical journals (1 general and 2 specialty) in The JAMA Network with the highest volume of research submissions in 2016: *JAMA* (n=4416), *JAMA Internal Medicine (JIM)* (n=1858), and *JAMA Pediatrics (JPED)* (n=1711). We compared rates of rejection without review vs peer review and acceptance vs rejection for papers with supplements vs papers without supplements and analyzed the peer reviewer comments for the supplements. In addition, we evaluated all research articles published in these 3 journals in 2016 (n=374) and the numbers and types of supplements published. For all articles published with supplements, we compared online usage (views and downloads) of articles vs the supplements for a defined period, January 2016-March 2017. For analysis of research submissions, we calculated risk differences with 95% CIs as well as 2-sided *P* values by χ^2 test.

Results Overall, 7985 research manuscripts were submitted to the 3 journals: 4868 (61%) with supplements and 3117 (39%) without supplements (**Table 60**). Papers with supplements were more likely to be peer reviewed than those without supplements: *JAMA* 34.1% vs 16.5% (difference, 17.6%; 95% CI, 15.1%-20.1%); *JIM* 17% vs 9% (difference, 8.0%; 95% CI, 4.8%-11.1%); *JPED* 32.9% vs 22.9% (difference, 10%; 95% CI, 5.8%-14.3%). Papers with supplements also more likely to be accepted than those without supplements: *JAMA* 3.9% vs 0.3% (difference, 3.6%; 95% CI, 2.9%-4.4%); *JIM* 8.4% vs 1.3% (difference, 7.1%; 95% CI, 5.4%-8.9%); *JPED* 8.7% vs 0.3% (difference, 8.4%; 95% CI, 6.6%-10.2%). Similar results were seen after excluding clinical trials and meta-analyses (for which the majority had supplements). Of 1421 papers with supplements that were sent for peer review, 484 (34%) were commented on by 538 reviewers. Of these comments, 155 (29%) were substantive comments, 63 (12%) indicated moving the supplemental content to the main paper, and 59 (11%) asked for more information about the supplement. Of 374 published research articles, 372 had supplements. Median online usage of articles far exceeded that of the supplements (*JAMA*: 10,918 vs 37 views or downloads; *JIM*: 3573 vs 12; *JPED*: 1386 vs 5). For articles, the majority of online usage (75%) occurred within 2 months of publication; for supplements, the majority of online usage (78%) took 7 months to occur.

Conclusions Although less valuable to peer reviewers and readers, online supplements appear to be a marker for complexity and quality of research manuscripts submitted and published in these 3 medical journals.

Conflict of Interest Disclosures: Other than affiliations with The JAMA Network journals included in the study and eJournalPress (provider of the manuscript submission and production systems used in the study), no other conflicts of interest are reported. Annette Flanagan is the Peer Review Congress Executive Director but was not involved in the review of or decision for this abstract.

Funding/Support: No external funding provided.

TUESDAY

Readability of Open Access and Non-Open Access Articles Reporting Research Studies in Primary Health Care: A Cross-sectional Study

Shuhei Ichikawa,¹ Kae Uetani,^{2,3} Yoshihito Goto,³ Takanori Fujita⁴

Objective Open access journals can enhance accessibility to the scientific articles. However, publishing in open access journals may not assure enough accessibility unless readers can understand what the article says. If articles in open access journals are more readable for nonprofessional readers, open access journals can enhance accessibility to contents of their articles. Therefore, we investigated whether articles in open access journals are more readable than journals with subscription models via a cross-sectional study.

Design We identified clinical trials and observational studies published in journals in primary health care field from April 2010 to December 2016 using PubMed. Journals with 20 or more target articles were extracted, and the most recent 20 articles were analyzed. Journals were divided into 3 groups: full open access, hybrid open access, and completely subscription. The primary outcome was readability in the Introduction section estimated by Gunning Fog score. Gunning Fog score ranges from 6 (level for sixth-grade student) to 17 (level for college graduates); lower score indicates that the text is more readable. Secondary outcomes were number of sentences, number of words, number of words that have 3 or more syllables, mean sentence length, and average syllables per words. Journal Impact Factor was set as a covariate because high-impact journals generally set a more limited article length, which would contribute to more efficient sentences. Country of journal publisher was set as a covariate because familiarity with the English language could affect readability of the articles. Bayesian estimation with Markov Chain Monte Carlo was used to estimate whether articles in full open access journals and hybrid open access journals were more readable than those in subscription journals.

Results One hundred forty articles were extracted from 7 journals: *Ann Fam Med*, *Fam Med*, *J Gen Intern Med*, *Br J Gen Pract*, *Fam Pract*, *J Am Board of Fam Med*, and *BMC Fam Pract*. Fog scores and other parameters in each group are shown in **Table 61**. Mean Fog scores in each type of journals was 19, which exceeded the college graduate level (17). Mean (SD) FOG scores in full open-access journals (19.0 [2.1]; estimates 0.69; 95% CI, -0.40 to 1.78) and hybrid open access journals (19.0 [2.5]; estimates, 0.12; 95% CI, -0.96 to

¹The JAMA Network, Chicago, IL, USA, annette.flanagan@jamanetwork.org; ²eJournalPress, Rockville, MD, USA

Table 60. Association of Supplement Inclusion With Likelihood of Research Manuscript Being Peer Reviewed or Accepted, Numbers of Supplements Commented on by Peer Reviewers, and Online Usage of Articles With Supplements vs Supplements Alone in 3 Medical Journals, 2016^a

	Total	No. (%)		Difference, % (95% CI)
		With Supplement	Without Supplement	
JAMA				
Research submissions in 2016	4416	2532 (57.3)	1884 (42.7)	
Peer reviewed	1174	863 (73.5)	311 (26.5)	
Rejected without review	3242	1669 (51.5)	1573 (48.5)	
Peer reviewed, %		34.1	16.5	17.6 (15.1-20.1)
Accepted ^b	104	99 (95.2)	5 (4.8)	
Rejected	4312	2433 (56.4)	1879 (43.6)	
Accepted, %		3.9	0.3	3.6 (2.9-4.4)
Peer reviewer comments on supplements	863	301 (34.9)		
Articles published in 2016	162	160 (98.8)	2 (1.3)	
Online usage, median (IQR)				
Per article		10,918 (1324-35,877)		
Per supplement		37 (1-96)		
JAMA Internal Medicine				
Research submissions in 2016	1858	1325 (71.3)	533 (28.7)	
Peer reviewed	273	225 (82.4)	48 (17.6)	
Rejected without review	1585	1100 (69.4)	485 (30.6)	
Peer reviewed, %		17	9	8.0 (4.8-11.1)
Accepted ^b	119	112 (94.1)	7 (5.9)	
Rejected	1739	1213 (69.8)	526 (30.2)	
Accepted, %		8.4	1.3	7.1 (5.4-8.9)
Peer reviewer comments on supplements	225	82 (36.4)		
Articles published in 2016 ^b	117	117 (100)	0	
Online usage, median (IQR)				
Per article		3573 (568-17,654)		
Per supplement		12 (1-53)		
JAMA Pediatrics				
Research submissions in 2016	1711	1011 (59.1)	700 (40.9)	
Peer reviewed	493	333 (67.5)	160 (32.5)	
Rejected without review	1218	678 (55.7)	540 (44.3)	
Peer reviewed, %		32.9	22.9	10.0 (5.8-14.3)
Accepted ^b	90	88 (97.8)	2 (2.2)	
Rejected	1621	923 (56.9)	698 (43.1)	
Accepted, %		8.7	0.3	8.4 (6.6-10.2)
Peer reviewer comments on supplements	333	101 (30.3)		
Articles published in 2016	95	95 (100)	0	
Online usage, median (IQR)				
Per article		1386 (407-5947)		
Per supplement		5 (1-13)		

Abbreviation: IQR, interquartile range.

^aOf 7985 total research submissions, 4868 (61%) had supplements and 3117 (39%) did not have supplements. Peer reviewers commented on 484 of 1421 (34%) research supplements sent for peer review for all 3 journals. Of 374 total research articles published, 372 had supplements and 2 did not have supplements.

^bThe numbers of articles accepted were recorded in February 2017 (total N = 313; JAMA 104, JAMA Internal Medicine 119, JAMA Pediatrics 90).

Table 61. Readability Scores and Other Parameters in Introduction Section of Articles in Each Type of Journal

Category	Mean (SD)		
	Completely Subscription ^a	Hybrid Open Access ^b	Full Open Access ^c
Articles, No.	40	60	40
FOG score	19.6 (2.4)	19.0 (2.5)	19.0 (2.1)
Sentences	15.7 (4.9)	18.0 (8.7)	24.2 (13.2)
Words	351.0 (107.0)	392.3 (165.0)	534.8 (256.8)
Word with 3 or more syllables	93.4 (34.9)	98.0 (47.7)	128.5 (57.1)
Sentence length	22.7 (3.1)	22.4 (3.5)	23.0 (3.7)
Syllables per words	1.88 (0.10)	1.86 (0.13)	1.85 (0.11)

^a Ann Fam Med and Fam Med

^b J Gen Intern Med, Br J Gen Pract, and Fam Pract

^c J Am Board of Fam Med, and BMC Fam Pract

1.23) were not significantly different from subscription journals (mean [SD], 19.6 [2.4]).

Conclusion We found no difference in readability among full open access, hybrid open access, and completely subscription journals.

¹Department of Education and Research in Family and Community Medicine, Mie University Graduate School of Medicine, Mie, Japan, jwellyfish@outlook.com, ichikawa-s@clin.medic.mie-u.ac.jp; ²Medical Communication Department, Statcom Co Ltd, Tokyo, Japan; ³Department of Health Informatics, Kyoto University School of Public Health, Kyoto, Japan; ⁴KEIO University Platform for Healthcare Industries, Tokyo, Japan

Conflict of Interest Disclosures: None reported.

Funding/Support: Mr Ichikawa received a grant for younger researcher from Mie University in 2016.

Role of the Funder/Sponsor: Mie University had no role for this research, except for funding.

Reporting Guidelines

MONDAY

Interventions to Improve Adherence to Reporting Guidelines: A Scoping Review

David Blanco de Tena-Dávila,¹ Jamie Kirkham,² Douglas G. Altman,³ David Moher,⁴ Isabelle Boutron,⁵ Erik Cobo⁶

Objective To investigate interventions aiming to improve adherence to reporting guidelines.

Design Ongoing scoping review of interventions aiming to improve adherence to reporting guidelines. The review follows the Joanna Briggs Institute scoping review methods manual. Since we want to map the literature and comprehensively summarize the existing evidence, we consider the scoping review methodology the most suitable approach. After searching MEDLINE, EMBASE, and the Cochrane Library databases, as well as Google Scholar, from January 1, 1996, to March 31, 2017, we identified 15 articles evaluating different actions to improve adherence to

reporting guidelines. The reference lists of these articles are still to be screened. Articles suggesting but not assessing interventions were collected but have not yet been analyzed. The interventions found so far were classified according different criteria, in relation to the target population (journal policies, authors, editors, or reviewers) or the research stage at which they are performed (design, conduct, reporting, or peer review). After completing the review, descriptive statistical analysis will be performed to summarize the effect of the evaluated interventions on adherence to reporting guidelines. Moreover, a comprehensive summary of the suggested interventions found will be presented.

Results Preliminary results show that the majority of the interventions found (11/15 [73%]) have been assessed in the last 6 years, showing that developing strategies to improve adherence to reporting guidelines is becoming a critical issue in health research. Most of the interventions (10/15 [67%]) target journal policies, including weak or strong endorsement of reporting guidelines, compulsory trial registration, or active implementation of reporting guidelines over the course of the editorial process. Others target authors (3/15 [20%]) or reviewers (2/15 [13%]). Few interventions (4/15 [27%]) were evaluated by a randomized trial, and none of these few interventions targeted journal policies. Although the effect of the interventions found varies greatly among studies, results of the review to date suggest that active implementation strategies of reporting adherence to guidelines over the course of the editorial process tend to improve completeness of reporting.

Conclusions Few interventions aiming to improve adherence to reporting guidelines have been assessed. Our preliminary results show that most editorial policies have been adopted without the previous best evidence. Further efforts should be taken to evaluate interventions by randomized trials. Moreover, journals should take action to encourage the use of reporting guidelines from early stages of research and also to actively implement guidelines in the editorial process.

¹Statistics and Operations Research Department, Barcelona Tech, Barcelona, Spain, david.blanco.tena@upc.edu; ²Department of Biostatistics, University of Liverpool, Liverpool, UK; ³Centre for Statistics in Medicine, University of Oxford, Oxford, UK; ⁴Ottawa Hospital Research Institute and University of Ottawa, Ottawa, Ontario, Canada; ⁵Centre d'Epidémiologie Clinique, Paris Descartes University, Paris, France; ⁶Statistics and Operations Research Department, Barcelona Tech, Barcelona, Spain

Conflict of Interest Disclosures: None reported. Douglas G. Altman, David Moher, and Isabelle Boutron are members of the Peer Review Congress Advisory Board but were not involved in the review of or decision for this abstract.

Funding/Support: This scoping review belongs to the ESR 14 research project from the Methods in Research on Research (MiRoR) project, which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 676207.

Role of Funder/Sponsor: The funders had no role in the design and conduct of this study.

A Qualitative Assessment of the STROBE Extensions: Laying the Groundwork for Future Educational Interventions

Melissa K. Sharp,^{1,2} Darko Hren²

Objective The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guideline was developed in response to inadequate reporting of observational studies. In recent years, several extensions to STROBE have been created to provide more nuanced field-specific guidance for authors. This evaluation aims to classify the changes made in the extensions in order to identify problem areas and the level of specificity needed for future educational interventions.

Design Two independent researchers assessed additions in each extension and achieved consensus; the intraclass correlation was calculated to measure agreement (ICC = 0.92). Individual additions were grouped by STROBE checklist items to identify the frequency and distribution of changes. Additions were coded as field specific or not field specific. *Field specific* was defined as information that is particularly relevant for a single field, and its guidance generally cannot be extrapolated outside that extension's field. *Not field specific* was defined as information that reflects a general epidemiological tenet and can be extrapolated to most if not all types of observational research studies.

Results A total of 297 additions were made across 13 STROBE extensions, with 36.7% of items classified as nonspecific (**Table 62**). The Methods section of STROBE contained the top 5 areas changed: statistical methods (44 additions; 45.5% nonspecific), participants (29 additions; 41.4% nonspecific), variables (28 additions; 32.1% nonspecific), setting (21 additions; 14.3% nonspecific), and study design (19 additions; 5.3% nonspecific). The items with the largest percentage of nonspecific recommendations were other additions (83.3%), bias (80.0%), other analyses (62.5%), and study size (60.0%). The quality of extensions was variable, with a range of 25% to 100% field-specific recommendations.

Conclusions One-third of all STROBE extension recommendations were not field specific, thus highlighting gaps in understanding of epidemiological principles or deficiencies in the scope or content of the original STROBE items. Next steps include a bibliometric study to establish the prevalence of extension endorsement as well as its effect on the completeness of reporting. Results from this study will determine the association between extension endorsement and completeness of reporting. This work will form a basis for a survey of author's knowledge, awareness, and use of STROBE to be distributed in early 2018. Together, these works will inform the creation of an educational intervention for authors reporting results from observational studies.

Table 62. STROBE Extension Additions

Section	STROBE Checklist Item	Total Additions, No.	Field-Specific Items, No. (%)
Title and abstract	Title and abstract	11	8 (72.7)
Introduction	Background or rationale	6	6 (100)
	Objectives	6	5 (83.3)
Methods	Study design	19	18 (94.7)
	Setting	21	18 (85.7)
	Participants	29	17 (58.6)
	Variables	28	19 (67.9)
	Data sources	19	11 (57.9)
	Bias	5	1 (20.0)
	Study size	5	2 (40.0)
	Quantitative variables	6	5 (83.3)
	Statistical methods	44	24 (54.5)
Results	Participants	18	14 (77.8)
	Descriptive data	17	11 (64.7)
	Outcome data	11	7 (63.6)
	Main results	16	7 (43.8)
	Other analyses	8	3 (37.5)
Discussion	Key results	0	0
	Limitations	11	5 (45.5)
	Interpretation	3	2 (66.7)
	Generalizability	2	2 (100)
Other	Funding	0	0
	Other additions	12	2 (16.7)

¹Department of Epidemiology, Paris Descartes University, Paris, France, msharp@unist.hr; ²Department of Psychology, School of Humanities and Social Sciences, University of Split, Split, Croatia

Conflict of Interest Disclosures: None reported.

Funding/Support: This work was supported by the European Union's Horizon 2020 research and innovation program under Marie Skłodowska-Curie grant agreement 676207.

MONDAY

Transparency and Completeness in the Reporting of Stakeholder Involvement in the Development and Reporting of Research Reporting Guidelines

Karen L. Woolley,¹⁻³ Serina Stretton,⁴ Lauri Arnsstein⁵

Objective The Guidance for Developers of Health Research Reporting Guidelines recommend multidisciplinary stakeholder involvement, transparent and complete reporting, and updating guidelines based on feedback. Developers are accountable for stakeholder engagement, but how broad and meaningful is such engagement? Our objective was to provide empirical feedback to developers by investigating (1) the involvement of those ultimately affected by guidelines (eg, patients and carers) and regular end users of guidelines (eg, publication professionals), and (2) the

transparency and completeness of reporting stakeholder involvement.

Design For this prospective study, conducted from September 2016 to January 2017, we included every reporting guideline for the main study types, as listed on the EQUATOR Network website. We pilot-tested a standardized data collection spreadsheet to extract data from the corresponding guideline publications. We quantified patient, carer, and publication professional involvement and used statisticians (listed as stakeholders in the guidelines) as a control group. We assessed reporting transparency and completeness using the AGREE reporting checklist for documenting stakeholder involvement. For qualitative insights, we interviewed leaders from nonprofit, international, patient advocacy (International Alliance of Patients' Organizations [IAPO]) and publication professional (Global Alliance of Publication Professionals [GAPP]) organizations.

Results Of the 33 guideline publications, the mean (SD) number of authors was 9 (5.7) (median, 7; IQR, 5-11) and the mean (SD) number of working group members was 45 (38.4) (median, 30; IQR, 23-43). Statisticians were authors of 8 publications (24%) and were working group members for 5 publications (15%). Patients, carers, and publication professionals were rarely identified, either as authors (0, 0, and 0, respectively) or working group members (0, 1 [3%], and 0, respectively). Reporting stakeholder involvement was deficient (eg, for statistician involvement, less than 25% publications met AGREE recommendations). Leaders from IAPO and GAPP were not aware of having been invited to participate in developing guidelines but thought that their stakeholders could provide unique and important insights. They encourage guideline developers to contact them to facilitate meaningful involvement.

Conclusions Guideline developers have rarely involved the stakeholders affected by guidelines (patients, carers) or those regularly using guidelines (publication professionals) in the development process. The involvement of these key stakeholders could enhance the credibility, dissemination, and use of guidelines. If patients, carers, and publication professionals were represented by other stakeholders (which is not ideal given potential conflicts of interest), this was not documented; readers do not know who represented whom. The transparency and completeness of reporting of stakeholder involvement must be improved.

¹ProScribe KK Envision Pharma Group, Tokyo, Japan; ²University of Queensland, Brisbane, Queensland, Australia, karen.woolley@envisionpharmagroup.com; ³University of the Sunshine Coast, Maroochydore DC, Queensland, Australia; ⁴ProScribe Envision Pharma Group, Sydney, New South Wales, Australia; ⁵Alligent Envision Pharma Group, London, UK

Conflict of Interest Disclosures: The authors acknowledge the valuable feedback provided by the International Alliance of Patients' Organizations during the development of the abstract. All authors are employees of the Envision Pharma Group and members of nonprofit associations that support ethical practices for publication professionals. Dr Woolley serves on the Patient Safety and Quality Committee of the Sunshine Coast Hospital and Health Services

Board, served on the Board of the International Society for Medical Publication Professionals (Asia Pacific Trustee), cofounded the Global Alliance of Publication Professionals (GAPP), and has shares in Johnson & Johnson. Dr Stretton is a voluntary member of GAPP. No other disclosures are reported.

MONDAY

Evaluation of Reporting Guideline Implementation by Editors of Rehabilitation-Related Journals

Allen Heinemann,¹ Leighton Chan,¹ Helen Hoenig,¹ Glenn Collins,¹ Jason Roberts²

Objective To describe the experience of rehabilitation journal editors in implementing reporting guidelines for original research articles.

Design We distributed, via email, an online survey to editors of 35 rehabilitation journals who agreed in 2013 to support the use of reporting guidelines. We emailed 2 reminders to complete the survey.

Results The response rate was 66% (23 of 35 editors). The editors received a mean of 435 submissions in 2016 (range, 61 to 1766) and published a mean of 93 manuscripts (range, 20 to 312). We reviewed author guidelines for the 12 journals whose editors did not respond to the survey and coded their guideline requirements. Overall, 60% (21 of 35) of journals require reporting guidelines. Editors of the 14 journals not requiring authors' use of reporting guidelines provided various reasons despite their 2013 support of reporting guidelines, including the belief that they did not improve manuscript quality. Among the 23 responding editors, CONSORT and PRISMA were required by 16 (70%), STROBE by 14 (61%), STARD by 13 (57%), and CARE by 9 (39%). Only 3 (13%) of journals allow exceptions. Most (12 of 14, 86%) require guideline checklist upload with submission. Only 6 (26%) of responding editors involve a statistical expert in reviewing every article. Time estimates to review guidelines on an individual article ranged from less than 15 minutes (7 of 16, 44%) to 15 to 30 minutes (7 of 16, 44%) to more than 30 minutes (2 of 16, 12%). While 10 of 17 responding editors (59%) believe authors accurately complete reporting guideline checklists with a few exceptions, the others (7 of 17, 41%) perceived quite a few exceptions to authors' accuracy and completeness. Most (12 of 21, 57%) believed that reporting guidelines resulted in a great deal of improvement in the quality of submitted manuscripts while others perceived some improvement (8 of 21, 38%) or no improvement (1 of 21, 5%). Editors identified authors' lack of familiarity with reporting guidelines as the largest barrier to reporting guidelines use; other barriers included increased time on editorial staff and costs associated with extra pages. Facilitators included mandatory use of guidelines and author educational efforts. Recommendations included promoting cooperation among editors to implement reporting guidelines, long lead times before requiring use of reporting guidelines, educating authors on the use of guidelines through links to outside

resources, and detailed author instructions. Editors valued the education resources provided by the EQUATOR Network.

Conclusions Editors of rehabilitation journals perceive the mandatory use of reporting guidelines to result in modest to considerable improvement in the quality of submitted manuscripts. They identified several strategies to promote author cooperation.

¹*Archives of Physical Medicine and Rehabilitation*, Reston, Virginia, USA, aheinemann@acrm.org; ²Origin Editorial, Leander, TX, USA

Conflict of Interest Disclosures: None reported.

Funding/Support: None reported.

TUESDAY

Journal Support for ARRIVE Guidelines and Reporting Quality of Animal Welfare, Analgesia, or Anesthesia Articles

Vivian Leung,¹ Frédéric Rousseau-Blass,¹ Daniel S. J. Pang¹

Objective The ARRIVE (Animal Research: Reporting of In Vivo Experiments) guidelines, published in 2010, were developed to improve the quality of animal research reporting. We hypothesised that articles published in veterinary journals supporting the ARRIVE guidelines would show improved quality of reporting compared with nonsupporting journals.

Design We identified veterinary journals that were likely to publish articles on animal welfare, analgesia, and anesthesia, topics of focus owing to their importance to animal well-being and potential influence on translational research. We defined animal welfare studies as those reporting interventions to improve animals’ environmental or physical conditions (eg, housing, enrichment). We distinguished journals that described the guidelines in their Instructions to Authors (guideline supporters [SUPP], n=5) from those that did not (nonsupporters [nonSUPP], n=2). Studies were identified by manual search of tables of contents (title, abstract, and keywords). The 20 items of the ARRIVE checklist were categorized by 2 independent authors (V.L. and F.R.B.) as fully, partially, or not reported, with differences resolved by consensus. We then compared adherence to guideline items in articles published pre-ARRIVE (2009) and post-ARRIVE (2015) in SUPP and nonSUPP journals using an unequal variance *t* test, and compared the difference in change.

Results A total of 236 papers were included: 120 from 2009 (SUPP, n = 52; nonSUPP, n = 68) and 116 from 2015 (SUPP, n = 61; nonSUPP, n = 55). There was no statistically significant difference between journal-type in the percentage of fully reported items in 2009 vs 2015 (Table 63). There were small, statistically significant increases in the percentage of reported items within journal type between 2009 and 2015, but no difference in the increase (absolute difference in change between NonSUPP and SUPP, 3.26%; 95% CI, -0.54% to 4.3%]; *P* = .09). No paper fully reported 100% of items on the ARRIVE checklist. Full reporting of several items

Table 63. Differences Between Journal Types and Years Before (2009) and After (2015) ARRIVE Guidelines Publication.

Variable	2009	2015	Change [95% CI] <i>P</i> Value ^b
SUPP, mean % (SD)	55.3 (11.5)	60.5 (11.2)	5.2 [1.0 to 9.4] .02
Non-SUPP, mean % (SD)	51.8 (9.0)	60.2 (10.0)	8.4 [5.0 to 11.8] >.001
Change, % [95% CI] <i>P</i> value ^a	3.5 [-0.3 to 7.3] .07	0.3 [-3.6 to 4.2] .89	NA

Values are mean percentages of fully reported items from papers included in each data set. ^a*P* values of differences between journal types within the same year. ^b*P* values of differences between years of the same journal type.

was low across journals and years: study design (< 30%), sample size justification (< 15%), allocation to experimental groups (< 30%), housing and husbandry details (< 20%), and experimental animals details (< 25%).

Conclusions Journal support of ARRIVE guidelines did not result in improved reporting in this sample. The standard of reporting was low, reflecting a need for animal journals to not only support but more actively enforce adherence to the ARRIVE guidelines. Our results are in agreement with previous studies assessing reporting standards pre-ARRIVE and post-ARRIVE publication.

¹Faculté de Médecine Vétérinaire, Université de Montréal, Saint-Hyacinthe, Québec, Canada, daniel.pang@umontreal.ca

Conflict of Interest Disclosures: None reported.

Funding/Support: DSJ Pang holds a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (ID: 424022-2013). F Rousseau-Blass is supported by the Fondation J-Louis Lévesque.

Role of the Funders: The funders had no role in study design, data collection and analysis, or decision to publish.

Reproducible Research

TUESDAY

Association of Random Audits of Researchers With Improved Overall Quality of Research

Adrian G. Barnett,¹ Nicholas Graves,¹ Pauline Zardo,²

Objective The “publish or perish” incentive drives researchers to focus on volume and novelty, which crowds the literature with unreliable findings. Different incentives are needed that prioritize quality. We examined whether random audits of researchers could improve quality.

Design We reproduced a simulated Darwinian model of research laboratories where the most successful laboratories produce “children” laboratories that inherit their parents’ characteristics (mutation probability, 0.01). Success is determined by the number of positive findings, which drives laboratories to publish multiple positive findings regardless of their veracity. Two key features of a laboratory are its power

to detect positive findings (fixed at 0.8) and the effort it exerts, which controls whether it starts a new hypothesis. We used the assumptions of the original simulation and extended the model to include random audits that examine laboratories' entire publication histories and calculated the false-positive rate with knowledge of which hypotheses are true. Laboratories with at least 50 publications with false-positive rates in the bottom third were removed. Audited laboratories increased their effort, as did their parents and existing children. We estimated the costs of auditing by assuming that 1 scientist per month was needed to review 10 publications. We used 500 simulations per scenario.

Results Without audits, effort declined, and the false-positive rate reached two-thirds in almost every simulation, meaning that most positive findings were wrong (**Figure 14**). Auditing 1.94% of all publications avoided the explosive increase in false positives in 95.0% of simulations. Audits decreased the volume of publications as laboratories worked longer on each question. Auditing 1.94% would cost an estimated \$169 per publication in US dollars (95% CI,

\$152-\$171). Adding measurement error of $\pm 20\%$ did not affect the efficacy of the audits.

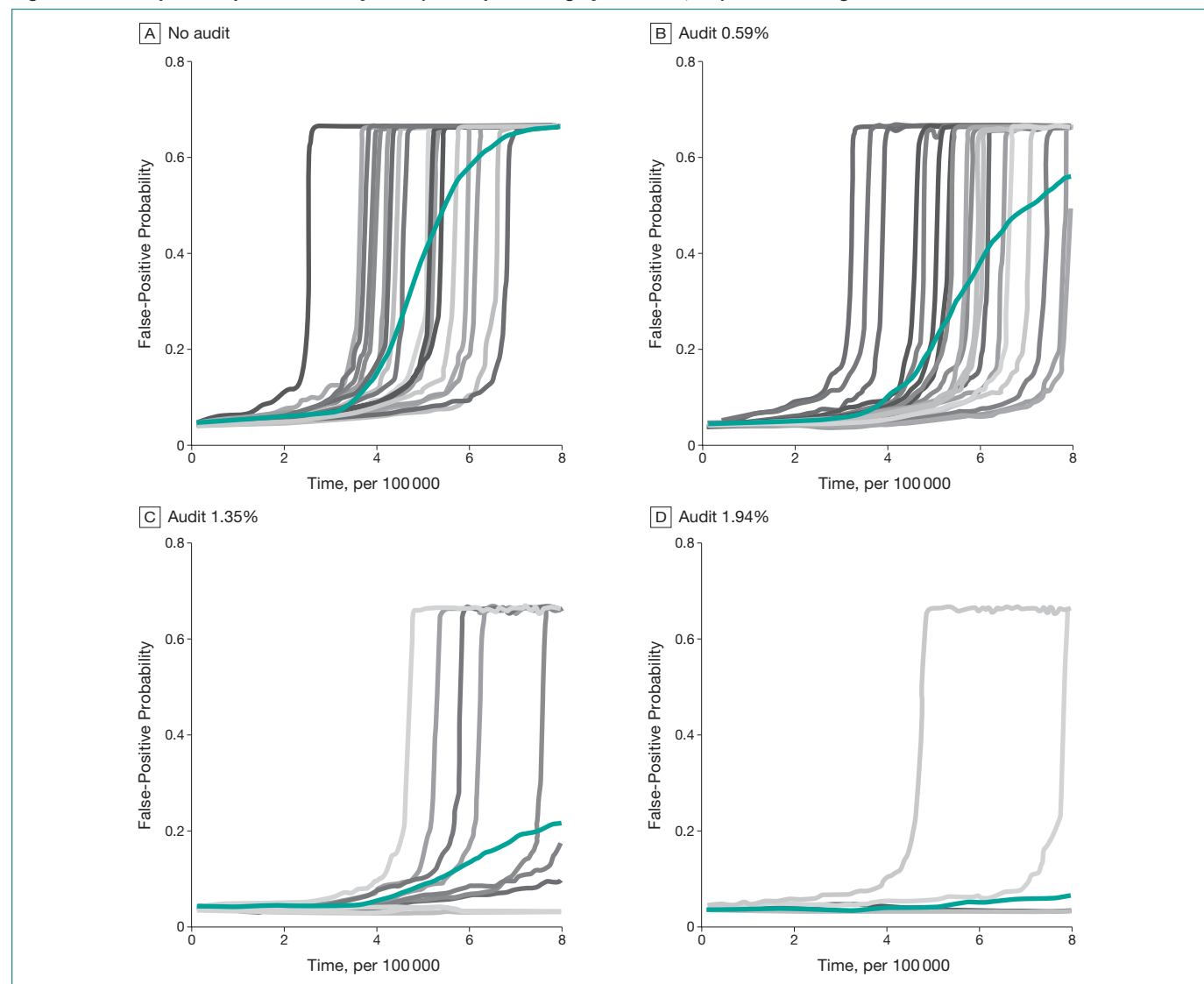
Conclusions Our simulation is a gross simplification of the research world, but the results provide food for thought. A random audit of a small proportion of the evidence base allows a detailed assessment of performance and may provide an incentive to researchers to raise standards. Audits are difficult to manipulate, but substantial funding and time would be needed from researchers to perform the audits. Audits could shift the maxim from "publish or perish" to "quality or perish."

¹School of Public Health and Social Work, Queensland University of Technology, Brisbane, Australia, a.barnett@qut.edu.au; ²Law School, Faculty of Law, Queensland University of Technology, Brisbane, Australia

Conflict of Interest Disclosures: None reported.

Funding/Support: Adrian G. Barnett is funded by a National Health and Medical Research Council (NHMRC) Senior Research

Figure 14. False-positive probabilities by time (No. of publishing cycles $\times 100,000$) for 4 auditing levels.



The gray lines are 20 randomly selected simulations; the dark red line in each graph is the average of 500 simulations.

Fellowship (APP117784). Nicholas Graves is also funded by an NHMRC Senior Research Fellowship (APP1059565).

Acknowledgments: Computational resources and services used in this work were provided by the High Performance Computer and Research Support Group, Queensland University of Technology, Brisbane, Australia.

Additional Information: For the statistical analysis, all the R code is available at <https://github.com/agbarnett/taxinspect>.

Research Methods

TUESDAY

Study Designs for the Evaluation of Biomarkers in Ovarian Cancer: A Systematic Review

Maria Olsen,¹ Mona Ghannad,^{1,2} and Patrick M. Bossuyt¹

Objective We documented the study design of recently reported evaluations of biomarkers in ovarian cancer to identify potential deficiencies.

Design We performed a systematic search in PubMed (MEDLINE) for reports of studies published in 2015 evaluating biomarkers in ovarian cancer using a combination of “ovarian cancer,” “biomarker,” and clinical performance measures and outcomes (eg, “survival,” “prognosis,” “prediction,” and “AUC”). Eligible were reports of clinical studies that had included adult patients; evaluated samples or data from women diagnosed, screened, treated, or monitored for ovarian cancer; and reported a clinical performance measure with no restrictions to study type (eg, diagnostic, predictive, and prognostic). We used the 1998 National Institutes of Health definition of a biomarker. Screening of titles and abstracts (level 1) and full text (level 2) was done in duplicate by 2 independent reviewers. Disagreements were solved through discussion, and a third reviewer was consulted if consensus was not reached. Data extraction from included study reports was done by 1 researcher and validated by a second. Using a structured data extraction form, we identified study design features and classified designs such as the use of single or multiple study groups, single or multicenter study, sample size, and characteristics of sample acquisition.

Results Our search resulted in 1026 studies: 516 (50%) and 345 (34%) were included from level 1 and 2, respectively, of which we evaluated the first 200 (starting January 1, 2015). We observed a wide diversity of study designs. In our preliminary results of 70 manuscripts, 36 (51%) were single group studies, whereas 27 (38%) included multiple groups (the remaining 8 (11%) being unclear); studies often used healthy controls 19 (27%); and 8 (11%) used extreme phenotypes. Fifty-nine (84%) had used pre-existing samples, while only 7 (10%) relied on dedicated acquisition of specimens and data. The sample size was limited in many studies.

Conclusions Our findings, show a variability in study designs for the evaluation of biomarkers in ovarian cancer and confirm the presence of suboptimal elements in recent evaluations of the clinical performance of biomarkers, such as

the inclusion of healthy control individuals, as well as limitations in the generalizability, with most studies being single-center. These limitations may contribute to failures in the translational phase of biomarker development, observed by other authors.

¹Academic Medical Center Amsterdam, Dept. of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam Public Health Research Institute, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands, m.olsen@amc.nl; ²Centre de Recherche Épidémiologie et Statistique Sorbonne Paris Cité, Université Paris Descartes, Centre d'épidémiologie Clinique, Hôpital Hôtel-Dieu, Paris, France

Conflict of Interest Disclosures: None reported.

Funding/Support: This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 676207.

Role of Funder/Sponsor: The funder played no role in this work.

Statistics

MONDAY

Benefits and Barriers to Implementation of Statistical Review at a Veterinary Medical Journal: A Mixed-Methods Study

Alexandra Winter,^{1,2} Nicola Di Girolamo,³ Michelle Giuffrida⁴

Objective To describe implementation of a statistical review process at a veterinary medical journal and perceived barriers to journal-wide adoption of this process.

Design In 2013 the *Journal of the American Veterinary Medical Association* assembled a group of experienced biostatistician reviewers, and 1 editor piloted a new statistical review process with the goal of achieving coordinated statistical review of all peer-reviewed manuscripts that included data. A collaborative approach was emphasized to maximize compliance. Other journal editors continued to request statistical review on an ad hoc basis, with biostatisticians made available to all editors starting in 2014. Peer review software was used to collect data on submission rates, numbers of statistical reviews, and turnaround times from editors and statistical reviewers. Qualitative data on perception of and barriers to implementation were collected from a convenience sample of authors, editors, and statistical reviewers.

Results Between 2013 and 2015, 2,198 manuscripts were submitted, including 1,021 primary research studies; 725 (71%) were sent for peer review, and 390 (38%) were accepted for publication. One hundred seven manuscripts underwent statistical review by 27 individuals (11 new statistical reviewers), totaling 166 statistical reviews (1-3 rounds). Ninety statistically reviewed primary research studies (84%) were accepted; of these, 61 (57%) were handled by the pilot editor, with the remainder handled by 6 other editors. For 2013-2015, the median time to final decision for primary research articles for the pilot editor was 242 days

(interquartile range, 65-527 days) vs 70 (interquartile range, 19-300) days overall. In qualitative responses, authors and statistical reviewers favored the process, despite increased time, effort, and communication required. Editors recognized the value added by the process but declined to endorse implementation of coordinated statistical review on a journal-wide basis. Perceived barriers to journal-wide adoption included increased editorial workload, resource limitations, increased turnaround time, and a belief that editors and standard peer reviewers could typically identify many statistical issues.

Conclusions The *Journal of the American Veterinary Medical Association* was able to implement a process of statistical review. Considerable communication by the pilot editor with authors and statistical reviewers facilitated implementation but also contributed to increased turnaround time, representing a logistical barrier. Authors indicated that statistical review improved their manuscripts. However, journal editors did not reach consensus on the need for a policy of journal-wide statistical review. These results illustrate the difficulty in balancing efforts to improve quality and transparency of reporting with expediency concerns and established editorial processes.

¹*Journal of the American Veterinary Medical Association*, Schaumburg, IL, USA, awinter@avma.org; ²American Veterinary Medical Association, Schaumburg, IL, USA; ³EBMVet, Cremona, Italy; ⁴Department of Surgical and Radiological Sciences, School of Veterinary Medicine, University of California-Davis, Davis, CA, USA

Conflict of Interest Disclosures: None reported.

Funding/Support: No external funding was provided for this study, which was conducted on a voluntary basis. Contribution of time and resources came from *the American Veterinary Medical Association*.

MONDAY

Authors' Assessment of the Impact and Value of Statistical Review in a General Medical Journal: 5-Year Survey Results

Catharine Stack,¹ Alicia Ludwig,² A. Russell Localio,³ Anne Meibohm,¹ Eliseo Guallar,⁴ John Wong,⁵ Deborah Cotton,¹ Cynthia Mulrow,¹ Jaya Rao,¹ Mary Beth Schaeffer,¹ Darren Taichman,¹ Christine Laine¹

Objective Statistical methods for biomedical research are increasingly complex. At our general medical journal, the statistical editors, in concert with the senior editors, perform statistical review of all provisional acceptances and revisions. Adding to our previously reported 2012 survey results, we sought authors' views annually for an additional 4 years regarding the impact of statistical review on the quality of their articles.

Design We conducted anonymous online annual surveys of corresponding authors of all articles published in 2012-2016 that underwent statistical review. We asked authors about the effort needed to respond to the statistical editors' requests and the impact of statistical review on the quality of both

specific sections and the overall published article. In 2014-2016 we also asked authors about the usefulness of optional conference calls with the editors, and an offer of such calls was included in all provisional acceptance letters sent during this period. The survey excluded rejected articles because they rarely receive full statistical review.

Results Of 489 authors (about 100 each year) surveyed, 337 (69%) responded. Response rates varied from 60% in 2016 to 79% in 2012. Studies included reports of original research (69%), systematic reviews and meta-analyses (24%), and decision analyses (7%). Fifty-seven percent of authors (range by year, 52%-61%) reported a moderate or large increase in the articles' overall quality as a result of the statistical editorial process; 54% noted improvements to the statistical methods section, 53% to the results section, and 32% to the conclusions section. Fewer authors reported no impact (15%) or a negative impact (2%) on the article. Fifty-eight percent of authors reported considerable effort to respond to the statistical editors' comments. A similar proportion (54%; range, 45%-65%) found that the effort required was worth the improved quality. Ten percent found the effort was not worth the improved quality. In 2014-2016, 19% (37/196) of respondents participated in conference calls with the statistical editors; 95% (35) found these calls extremely or somewhat helpful.

Conclusions In this survey conducted at a single, general medical journal, the majority of authors of articles published in the past 5 years reported improvements to their articles as a result of the statistical review and felt the effort required to respond to the statistical editors' comments was worth the improved quality. Telephone conferences were almost universally considered helpful.

¹*Annals of Internal Medicine*, American College of Physicians, Philadelphia, PA, USA, cstack@acponline.org; ²American College of Physicians, Philadelphia, PA, USA; ³University of Pennsylvania, Philadelphia, PA, USA; ⁴Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; ⁵Tufts University, Medford, MA, USA

Conflict of Interest Disclosures: Catharine Stack reports stock holdings in Pfizer and Johnson & Johnson. A. Russell Localio, Anne Meibohm, Eliseo Guallar, and John Wong report that funding from the *Annals of Internal Medicine* is provided to them or to their institution. Jaya Rao reports stock holdings/options in Eli Lilly and Pfizer. Christine Laine is a member of the Peer Review Congress Advisory Board but was not involved in the review of or decision for this abstract.

Funding/Support: No external funding was provided for this study. Contributions of staff time and resources came from *Annals of Internal Medicine*.

Trial Registration

TUESDAY

Proportion of National Institutes of Health R01-Funded Clinical Trials Registered in ClinicalTrials.gov

Erick H. Turner,^{1,2} An-Wen Chan,³ Dan A. Oren,⁴ Steven Bedrick¹

Objective Many clinical trials funded by the National Institutes of Health (NIH) are not published in a timely fashion. We sought to determine how many NIH-funded R01 clinical trials go unregistered.

Design We used the NIH RePORT database to identify R01 grants awarded from 2009 to 2013 for clinical trials by searching the (1) abstract field for “random%,” (2) the project terms field for “Randomized controlled trial(s),” “Randomized clinical trial(s) (RCTs),” “Randomized controlled clinical trial(s),” or “Randomized placebo controlled trial(s);” and (3) the NIH Spending Category field for “Clinical Trials.” We manually examined the context around term (1) and included only grants that clearly proposed new clinical trials; by the time of presentation, each full abstract will be rated by a second individual and rating discrepancies will be resolved by reaching a consensus. Using RePORT’s Clinical Studies tab, which links grants to ClinicalTrials.gov registrations, we identified trial grants that were linked to 1 or more registered trials. For grants lacking such links, we searched ClinicalTrials.gov by principal investigator last name and manually compared any matches with the index grant; we analyzed data for a random sample of 12 such unregistered grants but will report data for all “unregistered” clinical trial grants at the time of presentation.

Results Among 601 grants identified by search results, we excluded 116 after manual verification leaving 485 R01-funded clinical trials. The 485 grants were associated with 357 trial registrations. Excluding 24 redundant trials (those associated with the same grant) left 333 grants linked to 1 or more registered trials, suggesting a registration rate of 68.7%. However, among the random sample of 12 “unregistered” clinical trials, we were able to manually identify ClinicalTrials.gov registrations for 7.

Conclusion The proportion of NIH-R01-funded clinical trials registered at ClinicalTrials.gov is at least 68.7% and is likely higher. Reasons for nonregistration could include failure to conduct the trial after funding, which we could not measure. The RePORT database is a resource that can be used by systematic reviewers seeking comprehensive inception cohorts of NIH-funded clinical trials.

¹Oregon Health & Science University, Portland, OR, USA, turnere@ohsu.edu; ²Scientific Resource Center, Agency for Health Research Quality, Rockville, MD, USA; ³Women’s College Research Institute, Toronto, Ontario, Canada; ⁴Yale University, New Haven, CT, USA

Conflict of Interest Disclosures: Dr Turner receives protected time for research from the Scientific Resource Center, which is supported by the Agency for Health Research Quality. No other disclosures were reported.

Funding/Support: None reported.

Acknowledgments: The authors thank Stuart Buck, and Michael Stebbins, of the Arnold Foundation for discussions of the general idea of the study and Kevin Fain, Thiyagu Rajakannan, and Nicholas Ide, of ClinicalTrials.gov for discussions regarding link between ClinicalTrials.gov and RePORT.

MONDAY

Clinical Trials and Tribulations: “The Registration Haze”

Denise M. Goodman,¹ Karen E. Gutzman,² William F. Balistreri³

Objective As editors, we have encountered author confusion regarding which studies require trial registration and at what point in time this should be done. We hypothesized that this may be attributable to discrepancies between the instructions for authors of individual journals and requirements found on registry websites (eg, ClinicalTrials.gov) and on ICMJE.org and sought to describe these differences.

Design To be included, a journal had to have published at least 10 high-impact clinical trials from 2015 to 2016. A sample of more than 30 journals was identified for examination. Each journal title was searched using Google and the term <name of journal> instructions for authors to emulate author behavior. Journal websites were inspected for information on trial registration. Two investigators independently compared instructions for authors language with guidelines published at ClinicalTrials.gov and ICMJE.org and then coded the language as “conservative” (replicating ClinicalTrials.gov), “expanded” (replicating ICMJE.org), “other” (parochial, hybrid), or “silent” (no language provided). Discrepancies between investigators were resolved by discussion.

Results For the 32 journals inspected, the Impact Factor ranged from 10.95 to 59.56, and the journals published a median of 47.5 trials (range, 11-300). The κ value between the 2 investigators was 0.513, with disagreement on 10 of 32 journals. None of the instructions for authors were judged to be conservative, 15 as expanded, 11 as other, and 6 as silent. Of those coded as “other,” 1 journal differed from ICMJE in 2 aspects and all others in 1. Differences between ICMJE instructions and those from individual journals included timing of trial registration (before enrollment vs up to 6 months after, $n=1$), interventions requiring registration ($n=1$), and study design (requiring registration only for randomized clinical trials or a comparison or control group, $n=3$). The largest group ($n=7$) were nonspecific, stating only that the author should provide a registration number or listing approved registries. There was no association between code for registration instructions and number of citations or Impact Factor.

Conclusions More than one-third of high-impact journals did not follow either ICMJE or ClinicalTrials.gov guidelines, with one-fifth offering no guidelines on clinical trial registration. Harmonization of journal requirements with ICMJE could ameliorate this problem.

¹Northwestern University Feinberg School of Medicine, Ann & Robert H. Lurie Children’s Hospital of Chicago, Chicago, IL, USA, dgoodman@luriechildrens.org; ²Galter Health Sciences Library, Northwestern University Feinberg School of Medicine, Chicago, IL, USA; ³University of Cincinnati School of Medicine, Cincinnati, OH, USA

Conflict of Interest Disclosures: Denise Goodman is an associate editor at the *Journal of Pediatrics*, Elsevier, for which she receives an honorarium; the journal was not involved in this work. William Balistreri is editor in chief of the *Journal of Pediatrics*, for which he receives an honorarium; the journal was not involved in this work. Karen Gutzman reported no financial disclosures.

Funding/Support: This work was not funded.

TUESDAY

Adverse Event Reporting in Registered and Published Clinical Trials Focusing on Drug-Drug Interactions

Diana Jurić,¹ Shelly Pranić,² Ivančica Pavličević,³ Ana Marušić⁴

Objective Drug-drug interactions (DDIs) are a growing concern because of rising numbers of chemical entities and more prevalent polypharmacy. Several market withdrawals due to interaction-related adverse events (AEs) suggest the importance of reporting AEs in clinical studies focusing on DDIs. The aim of this study was to compare the characteristics of AEs in clinical trials focusing on DDIs registered in ClinicalTrials.gov and subsequently published in the medical literature.

Design As part of a larger study, we retrieved clinical trials focusing on DDIs from ClinicalTrials.gov using the search term *drug-drug interaction(s)* on October 16, 2015, and collected data on AEs from corresponding publications. Trials were included if they (1) primarily investigated DDIs; (2) had a ClinicalTrials.gov NCT number; (3) were registered between June 23, 2005, and October 16, 2015; and (4) were closed and completed by October 16, 2015. Published articles and their online supplementary material were identified from ClinicalTrials.gov and PubMed using the [si] tag along with the NCT number, and from SCOPUS/EMBASE using the first author's name and the study title. Data were abstracted by one author and verified by a second.

Results Among 2059 retrieved trials, 762 (37.0%) were excluded because they were not related to DDIs or had a change in the trial status. Of the remaining 1297 trials, most were interventional (1244 [95.9%]), industry funded (845 [65.2%]), and started before registration (745 [57.4%]). Trials were commonly phase 1 (895 [69.0%]), included healthy volunteers (873 [67.3%]), and had pharmacokinetic measures as the most common primary end point (633 [48.8%]). Results were registered for 164 trials (12.6%), among which 71 (43.3%) were published. However, the full text of 1 published trial was not available, leaving 70 trials in the analysis (Table 64). Published data about both serious and other AEs were identical to registered data for 17 trials (24.3%). Three trials (4.3%) with registered safety data did not describe them in publications. In 55 published trials with 1 or more other AEs recorded, equal absolute numbers and/or frequencies as in the registry were clearly reported for only 22 (31.4%). Different numbers of participants who discontinued treatment due to an AE were registered and reported for 2 trials.

Conclusions There are discrepancies between registered and published AE data for trials focusing on DDIs that emerge from incomplete or changed reporting of AEs in publications. There is a need to enforce regulatory requirements for timely and complete registration of results, and a need for clearer AE reporting for trials focusing on DDIs, including phase 1 trials, and for assessment of the congruence of registered and submitted AE data during the publication process.

¹Department of Pharmacology, University of Split School of Medicine, Split, Croatia, diana.juric26@gmail.com; ²Departments of Research in Biomedicine and Health and Public Health, University of Split School of Medicine, Split, Croatia; ³Department of Family Medicine, University of Split School of Medicine, Split, Croatia, ⁴Department of Research in Biomedicine and Health, University of Split School of Medicine, Split, Croatia

Conflict of Interest Disclosures: Dr Marušić is a Peer Review Congress Advisory Board Member but was not involved in the review or decision for this abstract.No other disclosures were reported.

Funding/Support: This study was funded by grant IP-2014-09-7672 ("Professionalism in Health Care") from the Croatian Science Foundation. The funder had no role in the design of this study or during its execution and data interpretation.

Table 64. Completeness of Registered and Published Safety Data in 71 ClinicalTrials.gov Trials on Drug-Drug Interactions^a

AEs in Registry vs Publications	No. (%) of Trials
SAEs = 0 in ClinicalTrials.gov	53 (75.7)
SAEs in publication	
Not reported	3 (4.3)
Reported as zero or not occurring	50 (71.4)
SAEs ≥1 in ClinicalTrials.gov	17 (24.3)
SAEs in publication	
Reported equal frequencies and/or absolute numbers	10 (14.3)
Reported less absolute numbers	1 (1.4)
Unclearly reported	6 (8.6)
OAEs = 0 in ClinicalTrials.gov	15 (21.4)
OAEs in publication	
Not reported	3 (4.3)
Reported as zero or not occurring	11 (15.7)
Reported OAE description without frequencies and/or absolute numbers	1 (1.4)
OAEs ≥1 in ClinicalTrials.gov	55 (78.6)
OAEs in publication	
Not reported	2 (2.9)
Reported equal frequencies and/or absolute numbers	22 (31.4)
Not comparable (unclearly reported or different reporting groups)	22 (31.4)
Different frequencies and/or absolute numbers	9 (12.9)
More in registry	3 (4.3)
More in publications	6 (8.6)

Abbreviations: AEs, adverse events; OAEs, other adverse events; SAEs, serious adverse events.
^aFull text of 1 publication was not available (n = 70 in analysis).

Adherence to the ICMJE Prospective Registration Policy Among Trials Published in High-Impact Specialty Society Journals

Anand D. Gopal,¹ Joshua D. Wallach,² Jenerius A. Aminawung,¹ Gregg Gonsalves,³ Rafael Dal-Ré,⁴ Jennifer E. Miller,^{5,6} Joseph S. Ross^{2,7}

Objective To evaluate adherence to the International Committee of Medical Journal Editors' (ICMJE) prospective registration policy, identify the frequency of registrations that occurred late enough to potentially permit protocol modifications based on premature examination of collected data, and determine characteristics associated with timely registration among clinical trials published in high-impact journals associated with US professional medical societies.

Design We conducted a cross-sectional analysis of the 50 most recently published clinical trials that reported primary results in the 10 highest-impact US specialty society journals between January 1, 2010 and December 31, 2015. We used descriptive statistics to characterize the proportions of clinical trials that were: registered on any of the 16 registries accepted by ICMJE; registered retrospectively; registered retrospectively potentially after initial ascertainment of primary outcomes; with concordant published and originally registered primary outcomes; and reporting favorable results, overall and stratified by journal and trial characteristics. χ^2 analyses were performed with a corrected type I error of 0.006 to assess differences in registration by journal and trial characteristics.

Results Among 6869 original research reports, we identified 472 articles reporting the primary results for 486 clinical trials. Of these 486 trials, 47 (10%) were unregistered, with proportions differing across journals. Among 439 registered clinical trials, 340 (77%) were registered prospectively and 99 (23%) retrospectively. Sixty-seven (15% of all registered trials) of these 99 retrospectively registered trials were registered late enough to have potentially permitted premature examination of primary outcome data ascertained among participants enrolled at inception. Among 413 clinical trials that registered and published at least 1 primary outcome, 109 (26%) published primary outcomes that differed from those first registered and 55 (13%) registered primary outcomes that were too poorly specified to permit comparison with published outcomes. Unregistered clinical trials were more likely to report favorable results than were registered clinical trials (89% vs. 64%; $P = .004$) irrespective of registration timing. FDA-regulated interventions, US-based studies, and industry funding were each associated with timely registration (**Table 65**).

Conclusions Adherence to ICMJE prospective registration policy remains sub-standard, even among the highest impact journals associated with US professional medical societies. These journals published unregistered trials and trials registered late enough to have potentially experienced

unaccounted protocol modifications after observation of primary outcomes.

¹Yale University School of Medicine, New Haven, CT, anand.gopal@yale.edu; ²Center for Outcomes Research and Evaluation, Yale-New Haven Hospital and Collaboration for Research Integrity and Transparency, Yale University, New Haven, CT; ³Yale School of Public Health, New Haven, CT; ⁴Epidemiology Unit, Health Research Institute-Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid, Madrid, Spain; ⁵Division of Medical Ethics, Department of Population Health, New York University School of Medicine, New York, NY; ⁶Bioethics International; ⁷Department of Internal Medicine, Yale University School of Medicine, New Haven, CT, USA

Conflict of Interest Disclosures: Mr Gopal receives research support through the Yale University School of Medicine Office of Student Research. Drs Wallach, Gonsalves, and Ross receive research support through Yale University from the Laura and John Arnold Foundation to support the Collaboration for Research Integrity and Transparency at Yale. Dr Miller also receives research support from the Laura and John Arnold Foundation. Dr Ross receives research support through Yale University from Johnson and Johnson to develop methods of clinical trial data sharing, from the Centers of Medicare and Medicaid Services to develop and maintain performance measures that are used for public reporting, from Medtronic, Inc. and the US Food and Drug Administration (FDA) to develop methods for postmarket surveillance of medical devices, from the FDA as part of the Centers for Excellence in Regulatory Science and Innovation (CERSI) program, and from the Blue Cross Blue Shield Association to better understand medical technology evaluation.

Funding Support: This project was not supported by any external grants or funds.

Acknowledgments: The authors would like to thank Alexander R. Bazazi, and Joseph Herrold, for their efforts in drafting earlier versions of the study protocol. Mr Bazazi and Dr Herrold were not compensated for their contributions to this work.

Table 65. Registration, Timeliness of Registration, Primary Outcome Concordance, and Primary Outcome Results Across Clinical Trials Published in the 10 Highest-Impact US Specialty Society Journals^a

Characteristic	No. (%)	Registration		Timeliness of Registration ^{b,c}				Primary Outcome Concordance ^d		Primary Outcome Results ^e	
		Unregistered (%) ^f	P Value	Retrospective (%) ^g	P Value	Retrospective after Initial Primary Outcome Ascertainment (%) ^{g,h}	P Value	Concordant (%) ^g	P Value	Favorable (%) ⁱ	P Value
Total	486 (100)	47 (9.7)		99 (22.6)		67 (15.3)		249 (56.7)		282 (66.4)	
Drug/device/ biological											
Yes	392 (80.7)	31 (7.9)	.008	65 (18.0)	<.001	42 (11.6)	<.001	216 (59.8)	.009	218 (64.7)	.16
No	94 (19.3)	16 (17.0)		34 (43.6)		25 (32.1)		33 (42.3)		64 (72.7)	
Funding ^j											
Industry	216 (44.4)	11 (5.1)	.002	25 (12.2)	<.001	18 (8.8)	<.001	131 (63.9)	.01	117 (65.7)	.82
Nonindustry	270 (55.6)	36 (13.3)		74 (31.6)		49 (20.9)		118 (50.4)		165 (66.8)	
Location											
≥1 US site	250 (51.4)	15 (6.0)	.005	35 (14.9)	<.001	23 (9.8)	<.001	133 (56.6)	.44	148 (65.8)	.79
Non-US	236 (48.6)	32 (13.5)		64 (31.4)		44 (21.6)		116 (56.9)		134 (67.0)	
Randomized											
Yes	372 (76.5)	23 (6.2)	<.001	79 (22.6)	.91	56 (16.0)	.39	195 (55.9)	.10	231 (64.0)	.01
No	114 (23.5)	24 (21.1)		20 (22.2)		11 (12.2)		54 (60)		51 (79.7)	
Enrollment											
≥ 100	280 (57.6)	9 (3.2)	<.001	58 (21.4)	.49	43 (15.9)	.69	162 (59.8)	.47	161 (62.7)	.05
< 100	206 (42.4)	38 (18.5)		41 (24.4)		24 (14.3)		87 (51.8)		121 (72.0)	

^aThe 10 highest-impact US specialty society journals were: *American Journal of Psychiatry*, *American Journal of Respiratory and Critical Care Medicine*, *Annals of Neurology*, *Blood*, *Gastroenterology*, *Hepatology*, *Journal of the American College of Cardiology*, *Journal of the American Society of Nephrology*, *Journal of Allergy and Clinical Immunology*, and *Journal of Clinical Oncology*.

^bClinical trials registered more than 30 days after enrollment started were considered to have been registered retrospectively. Note that ICMJE policy mandates registration before enrollment starts.

^cAmong 439 registered clinical trials, we could not determine the timeliness of registration for 2 (1 published in *Gastroenterology* and *Journal of Clinical Oncology*, respectively), as the enrollment start dates were missing from registrations. We excluded these 2 trials from analyses of association pertaining to overall timeliness of registration and timelines of registration relative to initial primary outcome ascertainment.

^dTwenty-six of 439 registered clinical trials did not have a primary outcome designated in their publication and were therefore excluded from analyses of association pertaining to primary outcome concordance.

^ePrimary outcome favorability could not be judged for 61 trials. These trials were excluded from analyses of association pertaining to primary outcome favorability.

^fPercentages are expressed as the fraction of total clinical trials in each row.

^gPercentages are expressed as the fraction of registered clinical trials (total-unregistered) in each row.

^hBecause of the nature of the primary outcome (ie, median survival), we could not determine if retrospective registration occurred after the initial primary outcome ascertainment in 8 cases: 1 in *Blood*; 1 in *Hepatology*; 2 in *Journal of Allergy and Clinical Immunology*; and 4 in *Journal of Clinical Oncology*. These trials were excluded from analyses of association pertaining to timeliness of registration relative to initial primary outcome ascertainment.

ⁱPercentages are expressed as the fraction of trials in each row for which primary outcome favorability could be judged (row totals not shown).

^jClinical trials receiving either full or partial industry support were designated as having received industry funding.

SPONSORS

GOLD

Meta-Research Innovation Center at Stanford (METRICS)

metrics.stanford.edu

The Meta-Research Innovation Center at Stanford (METRICS) is a research to action center focused on transforming research practices to improve the quality of scientific studies in biomedicine and beyond. METRICS fosters multi-disciplinary research collaborations to help produce solutions that increase the effectiveness and value of scientific investigation.

Wolters Kluwer Health

wolterskluwerhealth.com

Wolters Kluwer Health is a leading global provider of information and point of care solutions for the healthcare industry. Our solutions are designed to help professionals build clinical competency and improve practice so that healthcare organizations can succeed in value-based care delivery models. Product solutions include Lippincott, Ovid, and UpToDate.

SILVER

New England Journal of Medicine

www.nejm.org

The *New England Journal of Medicine* (NEJM.org) is the world's leading medical journal and website. *NEJM* publishes peer-reviewed research and interactive clinical content for physicians, educators, and the global medical community.

eJournalPress

www.ejournalpress.com

eJournalPress is focused on providing individualized online peer review, journal production tracking, and billing payment solutions to the scientific publishing community. Our goal is to find the best solutions for each organization's needs through innovative technology. Our Continuous Innovation Process allows us to tailor the software to fit any workflow.

BRONZE

ACP/Annals of Internal Medicine

www.acponline.org

Annals of Internal Medicine, celebrating its 90th year and known for excellence in peer review, is published by the American College of Physicians. *Annals* promotes excellence, advances standards in research methods, and contributes to improving health worldwide by publishing research, reviews, guidelines, policy papers, and commentaries relevant to internal medicine and its subspecialties.

Aries Systems Corporation

www.ariessys.com

Aries Systems transforms the way scholarly publishers bring high-value content to the world. The company's innovative and forward-looking workflow solutions help manage the complexities of modern print and electronic publishing—from submission, to editing and peer review, to production tracking. Publish faster, publish smarter, with Aries Systems.

BioMed Central, part of Springer Nature

www.biomedcentral.com

Born of innovation, we proved open access and open research as a sustainable model. We pride ourselves on continually improving our author services and the quality of our publishing. In 2016 BMC published over 40,000 open peer review reports in 70 journals – allowing 24,000 peer reviewers to be recognized for their contribution and helping younger scientists to learn from the work of others, giving peer reviewers a better and more transparent experience. Now as BMC and as part of Springer Nature, we extend this commitment to communities and disciplines beyond biology and medicine.

PLOS: Public Library of Science

www.plos.org

PLOS (Public Library of Science) is a nonprofit open access publisher, innovator and advocacy organization dedicated to accelerating progress in science and medicine by leading a transformation in research communication. The PLOS suite of influential journals contain rigorously peer-reviewed open access research articles from all areas of science and medicine.

PRE

www.pre-val.org

Peer-Review Evaluation (PRE) is a transparency service from AAAS that showcases the strength of your peer review process. PRE integrates with your existing workflows and technology platforms to analyze your manuscript tracking system data and tag qualifying articles with a PRE award badge to show readers which articles have been peer reviewed. For more information visit www.pre-val.org.

Silverchair Information Systems

www.silverchair.com

Silverchair integrates and delivers scholarly and professional content from a single platform — journals, books, video, custom formats, and more. The Silverchair Platform provides distinctive online sites, unique products, and advanced semantic technologies to STM and humanities publishers, professional societies, and the federal government, propelling their content to greater reach and impact.

Additional Sponsors

Copyright Clearance Center, Inc.

www.copyright.com

Copyright Clearance Center (CCC) is a global leader in content management, licensing, discovery and delivery solutions. Through its relationships with those who use and create content, CCC drives market-based solutions that accelerate knowledge, power publishing and advance copyright. With its subsidiaries RightsDirect and Ixxus, CCC provides solutions for millions of people from the world's largest companies and academic institutions.

HighWire Press, Inc.

www.highwire.org

HighWire Press, Inc. delivers technology solutions for leading publishers to disseminate scholarly research worldwide. With offices in the US and the UK, HighWire provides strategic consulting, professional services and innovative open platform technologies.

EXHIBITORS

Altmetric

www.altmetric.com

Who's talking about your research and how does this impact your organisation? Altmetric works with many of the world's leading funders, publishers, institutions, non-profit organisations and pharmaceutical companies to track the online engagement surrounding publications in real time, enabling users to monitor trends, identify engaged audiences and key opinion leaders, gather commercial intelligence and inform strategic decision making.

Aries Systems Corporation

www.ariessys.com

Aries Systems transforms the way scholarly publishers bring high-value content to the world. The company's innovative and forward-looking workflow solutions help manage the complexities of modern print and electronic publishing—from submission, to editing and peer review, to production tracking. Publish faster, publish smarter, with Aries Systems.

Cenveo Publisher Services

www.cenveopublisherservices.com

Cenveo Publisher Services is the industry leader in transformative publishing solutions. Cenveo supports commercial and scholarly journal publishers with customized peer review management services tailored to all journals and requirements. For more information, please visit www.cenveopublisherservices.com.

Clarivate Analytics

clarivate.com

Clarivate Analytics accelerates the pace of innovation by providing trusted insights and analytics to customers around the world, enabling them to discover, protect and commercialize new ideas faster. Formerly the Intellectual Property and Science business of Thomson Reuters, we own and operate a collection of leading subscription-based services focused on scientific and academic research, patent analytics and regulatory standards, pharmaceutical and biotech intelligence, trademark protection, domain brand protection and intellectual property management. Clarivate Analytics is now an independent company with over 4,000 employees, operating in more than 100 countries with well-known brands including Web of Science, Cortellis, Thomson Innovation, Derwent World Patents Index, CompuMark, MarkMonitor, Techstreet, and others.

Delta Think, Inc.

deltathink.com

Delta Think is a consultancy and advisory firm focused on strategy and innovation in scholarly communications and association management. We help organizations clarify business objectives, gain customer insight, define digital strategies, and implement new products and technologies. In 2016, Delta Think launched the Open Access Data Analytics Tool to support open access publishers making data-driven decisions.

Editage/Cactus Communications

www.editage.com

Founded in 2002, Editage, Cactus Communications, is a scholarly communications company with 640+ employees and offices in Japan, South Korea, India, China, and the US. Editage has served 183,000 authors across 1,200 disciplines, and partners with publishers, journals, academic societies, and pharma worldwide. Editage also offers comprehensive author education through Editage Insights.

eJournalPress

www.ejournalpress.com

eJournalPress is focused on providing individualized online peer review, journal production tracking, and billing payment solutions to the scientific publishing community. Our goal is to find the best solutions for each organization's needs through innovative technology. Our Continuous Innovation Process allows us to tailor the software to fit any workflow.

EQUATOR Network

www.equator-network.org

The EQUATOR Network is an international initiative dedicated to improving the reliability of the published literature. We promote the use of reporting guidelines to help researchers plan and write up their work clearly, completely and accurately so it's reproducible and useful to readers of journal articles.

HighWire Press, Inc.

www.highwire.org

HighWire Press, Inc. delivers technology solutions for leading publishers to disseminate scholarly research worldwide. With offices in the US and the UK, HighWire provides strategic consulting, professional services and innovative open platform technologies.

Inera

www.inera.com

Inera is the developer of the eXtyle family of Word-based editorial and XML tools and the Edifix online bibliographic reference solution. Publishers of scholarly journals and books, standards, and government documents worldwide rely on Inera's solutions to drive modern electronic and print publication workflows. Inera actively engages in standards development (NLM/JATS/BITS DTDs), cross-publisher initiatives (Crossref), and board membership in key organizations (NISO).

J&J Editorial

jjeditorial.com

J&J Editorial provides customized editorial and production solutions to editorial offices, publishers, and societies. J&J's employees work with clients to understand their unique needs and offer a range of editorial, production, copyediting, and submission systems support services to serve clients at a consistent level of excellence.

Origin Editorial

origineditorial.com

Origin Editorial provides efficient, evidence-based editorial office management, and consultation on improving efficiency in the editorial office and negotiating contracts with publishers. We focus on professional communication and proactive management of the peer review process, anticipating, and not just reacting to, problems. Using our knowledge of best practices, we will unlock your journal's potential.

Overleaf

www.overleaf.com

Founded in 2012 and with over 600,000 registered users, Overleaf is an academic authorship tool that allows seamless collaboration and effortless manuscript submission, all underpinned by cloud-technology. By providing an intuitive online collaborative writing and publishing platform, Overleaf is making the process of writing, editing and publishing scientific documents quicker and easier. Researchers and academics can now write, collaborate, and publish with a single click, directly from the Overleaf web-app. Publishers and institutions are partnering with Overleaf to provide customized writing templates, simple reference tool linking, and one-click publishing submission links.

Oxford University Press

global.oup.com

Oxford University Press publishes some of most respected and prestigious books and journals in the world. Visit our booth or www.global.oup.com for more information.

PleaseTech Ltd

www.pleasetech.com

PleaseTech specializes in software solutions for the collaborative review, co-authoring and redaction solution of documents. Our flagship product, PleaseReview, solves the critical business issue of document review by offering controlled, simultaneous and secure collaboration for the review/editing of documents. PleaseReview is used extensively by medical writers and medical writing consultancies.

Technica Editorial Services

technicaeditorial.com

Technica Editorial brings nearly three decades of practical publishing and editorial experience to its clients, which include commercial and not-for-profit publishers, associations, and societies. From basic peer review support to total project management, Technica tailors its services to satisfy each project's unique requirements. Technica Editorial — Proof that support matters.

Eighth International Congress on Peer Review and Scientific Publication Organizers and Planners

The JAMA Network

www.jamanetwork.com

Building on a tradition of editorial excellence, The JAMA Network brings *JAMA* together with the 11 *JAMA* specialty journals to offer enhanced access to the research, reviews, and perspectives shaping the future of medicine. Through a variety of innovative tools, The JAMA Network provides the insights that matter most to medical research and practice. The journals are published in print, online, and via The JAMA Network Reader. The JAMA Network Journals generate about 75 million article views per year.

BMJ

www.bmj.com/company

BMJ advances healthcare worldwide by sharing knowledge and expertise to improve experiences and outcomes. We publish one of the world's most cited general medical journals, *The BMJ*, and over 60 specialty journals. We also offer digital professional development courses and clinical decision support tools to help healthcare professionals improve the quality of healthcare delivery. We are owned by the British Medical Association (BMA), the professional association for doctors in the UK.

Meta-Research Innovation Center at Stanford (METRICS)

metrics.stanford.edu

The Meta-Research Innovation Center at Stanford (METRICS) is a research to action center focused on transforming research practices to improve the quality of scientific studies in biomedicine and beyond. METRICS fosters multi-disciplinary research collaborations to help produce solutions that increase the effectiveness and value of scientific investigation.

Thank you!

The organizers wish to thank all plenary session and poster session presenters, moderators, advisory board members, sponsors, exhibitors, and all participants, and the following individuals who contributed significantly to the planning and support of this Congress:

Jeni Reiling, Rosa Miranda, Caroline Sietmann, Connie Murphy, Connie Manno, Joe Amft, David Antos, Rick Bell, Deanna Bellandi, Karen Boyd, Kevin Brown, Michael Bruny-Groth, Larry Bryant, Eric Butkus, Carrie Butt, Debra Camp, Kim Campbell, Wendy Carranza, Lindsay Dawson, Gabriel Dietz, Amanda Ehrhardt, Nancy Essex, Carol Joy Farrell, Maria Ferrara, Ted Grudzinski, Mustafa Ishaq, Millette Jackson-Bates, Joy Jaeger, Erin Kato, Emily Ling, Rachell Lozano, Craig McCaffrey, John McFadden, Chris Meyer, Jim Michalski, Reuben Rios, Paul Ruich, Michael Ryder, Phil Sefton, Shannon Sparenga, Beverly Stewart, Suzanne Walker, Kim Westman, Sam Wilder, and Elizabeth Zak.



With Sponsorship From

GOLD

Meta-Research Innovation Center at Stanford (METRICS)

metrics.stanford.edu

Wolters Kluwer Health

wolterskluwerhealth.com

SILVER

New England Journal of Medicine

www.nejm.org

eJournalPress

www.ejournalpress.com

BRONZE

ACP/Annals of Internal Medicine

www.acponline.org

Aries Systems Corporation

www.ariessys.com

BioMed Central, part of Springer Nature

www.biomedcentral.com

PLOS: Public Library of Science

www.plos.org

PRE

www.pre-val.org

Silverchair Information Systems

www.silverchair.com

Additional Sponsors

Copyright Clearance Center, Inc.

www.copyright.com

HighWire Press, Inc.

www.highwire.org

